

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 11, 2004, 10:08:26 ; Search time 214 Seconds
(without alignments)

3127.023 Million cell updates/sec

Title: US-10-009-317A-30

Perfect score: 132

Sequence: 1 ggacctgtctactagtcc.....cggcaacaataattacatg 132

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:
1: /cgn2_6/ptodata1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata1/pubpna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 132 | 100.0 | 132 | 15 | US-10-313-994-30 |
| 2 | 109.2 | 82.7 | 132 | 15 | US-10-313-994-31 |
| 3 | 30.6 | 23.2 | 450 | 13 | US-10-027-632-77713 |
| 4 | 30.6 | 23.2 | 450 | 16 | US-10-027-632-77713 |
| 5 | 30.6 | 23.2 | 521 | 15 | US-10-029-386-4727 |
| 6 | 30.6 | 23.2 | 1985 | 10 | US-09-933-767-97 |
| 7 | 30.6 | 23.2 | 1985 | 13 | US-10-004-860-97 |
| 8 | 30.6 | 23.2 | 1985 | 15 | US-10-023-282-97 |
| 9 | 30.6 | 23.2 | 1987 | 16 | US-10-264-237-115 |
| 10 | 30.2 | 22.9 | 2301 | 16 | US-10-369-493-33481 |
| 11 | 27.8 | 21.1 | 2000 | 9 | US-09-887-576-863 |
| 12 | 27.8 | 21.1 | 4565 | 17 | US-10-437-963-37892 |
| 13 | 27.2 | 20.6 | 35548 | 13 | US-10-087-192-1786 |
| 14 | 26.8 | 20.3 | 409 | 13 | US-10-027-632-38125 |

| | | | | | | | |
|---|----|------|------|---------|----|----------------------|--------------------|
| C | 15 | 26.8 | 20.3 | 409 | 16 | US-10-027-632-38125 | Sequence 38125, A |
| | 16 | 26.8 | 20.3 | 473 | 17 | US-10-767-701-17641 | Sequence 17641, A |
| | 17 | 26.8 | 20.3 | 62705 | 17 | US-10-317-803-4 | Sequence 4, Appl |
| C | 18 | 26.6 | 20.2 | 1875 | 9 | US-09-938-842A-694 | Sequence 694, App |
| | 19 | 26.6 | 20.2 | 1875 | 11 | US-09-938-842A-694 | Sequence 694, App |
| C | 20 | 26.6 | 20.2 | 3365 | 16 | US-10-108-260A-1417 | Sequence 1417, Ap |
| | 21 | 26.6 | 20.0 | 240825 | 10 | US-09-790-289-1 | Sequence 1, Appl |
| C | 22 | 26.4 | 20.0 | 240825 | 17 | US-10-468-582-1 | Sequence 1, Appl |
| | 23 | 26.2 | 19.8 | 577 | 13 | US-10-027-632-77120 | Sequence 77120, A |
| | 24 | 26.2 | 19.8 | 577 | 13 | US-10-027-632-77121 | Sequence 77121, A |
| | 25 | 26.2 | 19.8 | 577 | 13 | US-10-027-632-77122 | Sequence 77122, A |
| | 26 | 26.2 | 19.8 | 577 | 13 | US-10-027-632-300287 | Sequence 300287, A |
| | 27 | 26.2 | 19.8 | 577 | 13 | US-10-027-632-300288 | Sequence 300288, A |
| | 28 | 26.2 | 19.8 | 577 | 16 | US-10-027-632-77120 | Sequence 77120, A |
| | 29 | 26.2 | 19.8 | 577 | 16 | US-10-027-632-77121 | Sequence 77121, A |
| | 30 | 26.2 | 19.8 | 577 | 16 | US-10-027-632-77122 | Sequence 77122, A |
| | 31 | 26.2 | 19.8 | 577 | 16 | US-10-027-632-300287 | Sequence 300287, A |
| | 32 | 26.2 | 19.8 | 577 | 16 | US-10-027-632-300288 | Sequence 300288, A |
| | 33 | 26.2 | 19.8 | 1046 | 13 | US-10-027-632-300289 | Sequence 300289, A |
| | 34 | 26.2 | 19.8 | 1046 | 16 | US-10-027-632-300289 | Sequence 300289, A |
| C | 35 | 26.2 | 19.8 | 3099 | 13 | US-10-027-632-111711 | Sequence 111711, A |
| | 36 | 26.2 | 19.8 | 3099 | 16 | US-10-027-632-111711 | Sequence 111711, A |
| | 37 | 26.2 | 19.8 | 3100 | 13 | US-10-027-632-116401 | Sequence 116401, A |
| | 38 | 26.2 | 19.8 | 3100 | 16 | US-10-027-632-116401 | Sequence 116401, A |
| | 39 | 26.2 | 19.8 | 4161 | 16 | US-10-398-221-2255 | Sequence 2255, Ap |
| | 40 | 26.2 | 19.8 | 4185 | 16 | US-10-398-221-684 | Sequence 684, App |
| C | 41 | 26.2 | 19.8 | 10236 | 15 | US-10-114-170-240 | Sequence 240, App |
| | 42 | 26.2 | 19.8 | 786431 | 15 | US-10-412-277-3 | Sequence 3, Appl |
| C | 43 | 26.2 | 19.8 | 1163020 | 16 | US-10-398-221-10 | Sequence 10, Appl |
| | 44 | 26.2 | 19.8 | 3011208 | 16 | US-10-398-221-2058 | Sequence 2058, Ap |
| | 45 | 26 | 19.7 | 525 | 15 | US-10-081-051-61 | Sequence 61, Appl |

ALIGNMENTS

RESULT 1

US-10-313-994-30
; Sequence 30, Application US/10313994
; Publication No. US20030162718A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/10/313,994
; CURRENT FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 30
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-10-313-994-30

Query Match 100.0%; Score 132; DB 15; Length 132;
Best Local Similarity 100.0%; Pred. No. 7.8e-41;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|----|---|-----|
| Qy | 1 | GGACCTGTCTACTAGTCCCTTACCGAGTAGGAAGAAATGATACCGACCTTTTGGCGGAGAA | 60 |
| Db | 1 | GGACCTGTCTACTAGTCCCTTACCGAGTAGGAAGAAATGATACCGACCTTTTGGCGGAGAA | 60 |
| Qy | 61 | GGTGAAGTCTCAGTCGCTTTCCGAACTCCACGTAAACGTGTGCTCTTAAGACGGCAAC | 120 |
| Db | 61 | GGTGAAGTCTCAGTCGCTTTCCGAACTCCACGTAAACGTGTGCTCTTAAGACGGCAAC | 120 |

Db 228 ATTTC 232

RESULT 5
US-10-029-386-4727
; Sequence 4727, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annotax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4727
; LENGTH: 521
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AB022537.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.71
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.5
; OTHER INFORMATION: EST_HUMAN HIT: BG680371.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q61187, EVALUATE 9.60e-01
; OTHER INFORMATION: NT HIT: g114782799, EVALUATE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BG680371.1, EVALUATE 0.00e+00
US-10-029-386-4727

Query Match 23.2%; Score 30.6; DB 15; Length 521;
Best Local Similarity 52.8%; Pred. No. 0.29; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 59

QY 4 CCTGTCTACTAGTCCTCCCTTACCAGGTAAGGAAATGACCGGACTTTTGGCGCGAGAAGGT 63
Db 313 CCITCACTCCAGCAGCTTGCCCAACAGGATAAGCTGGATCCCTTGGCCTTCTGAATAT 372

QY 64 GAAAGTCTCAGTCGCTTTCCGAAGTCCAGTAAAGCTGTCTCTTAAGACGGCAAAACAT 123
Db 373 CCAGTGTCTCAGGTTTCCCAAGACCACTTCCCTGTGGGCTTCCAAAATGGCCTTTATC 432

QY 124 ATTAC 128
Db 433 ATTTC 437

RESULT 6
US-09-933-767-97
; Sequence 97, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P2
; CURRENT APPLICATION NUMBER: US/09/933,767
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: PCT/US01/05614
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,836
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/193,170
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/205,258
; PRIOR FILING DATE: 1998-12-04
; PRIOR APPLICATION NUMBER: PCT/US98/11422
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/048,885
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,375

| | |
|----|--------------------------------------|
| 1 | PRIOR FILING DATE: 1997-06-06 |
| 2 | PRIOR APPLICATION NUMBER: 60/048,881 |
| 3 | PRIOR FILING DATE: 1997-06-06 |
| 4 | PRIOR APPLICATION NUMBER: 60/048,880 |
| 5 | PRIOR FILING DATE: 1997-06-06 |
| 6 | PRIOR APPLICATION NUMBER: 60/048,896 |
| 7 | PRIOR FILING DATE: 1997-06-06 |
| 8 | PRIOR APPLICATION NUMBER: 60/049,020 |
| 9 | PRIOR FILING DATE: 1997-06-06 |
| 10 | PRIOR APPLICATION NUMBER: 60/048,876 |
| 11 | PRIOR FILING DATE: 1997-06-06 |
| 12 | PRIOR APPLICATION NUMBER: 60/048,895 |
| 13 | PRIOR FILING DATE: 1997-06-06 |
| 14 | PRIOR APPLICATION NUMBER: 60/048,884 |
| 15 | PRIOR FILING DATE: 1997-06-06 |
| 16 | PRIOR APPLICATION NUMBER: 60/048,894 |
| 17 | PRIOR FILING DATE: 1997-06-06 |
| 18 | PRIOR APPLICATION NUMBER: 60/048,971 |
| 19 | PRIOR FILING DATE: 1997-06-06 |
| 20 | PRIOR APPLICATION NUMBER: 60/048,964 |
| 21 | PRIOR FILING DATE: 1997-06-06 |
| 22 | PRIOR APPLICATION NUMBER: 60/048,882 |
| 23 | PRIOR FILING DATE: 1997-06-06 |
| 24 | PRIOR APPLICATION NUMBER: 60/048,899 |
| 25 | PRIOR FILING DATE: 1997-06-06 |
| 26 | PRIOR APPLICATION NUMBER: 60/048,893 |
| 27 | PRIOR FILING DATE: 1997-06-06 |
| 28 | PRIOR APPLICATION NUMBER: 60/048,900 |
| 29 | PRIOR FILING DATE: 1997-06-06 |
| 30 | PRIOR APPLICATION NUMBER: 60/048,901 |
| 31 | PRIOR FILING DATE: 1997-06-06 |
| 32 | PRIOR APPLICATION NUMBER: 60/048,892 |
| 33 | PRIOR FILING DATE: 1997-06-06 |
| 34 | PRIOR APPLICATION NUMBER: 60/048,915 |
| 35 | PRIOR FILING DATE: 1997-06-06 |
| 36 | PRIOR APPLICATION NUMBER: 60/049,019 |
| 37 | PRIOR FILING DATE: 1997-06-06 |
| 38 | PRIOR APPLICATION NUMBER: 60/048,970 |
| 39 | PRIOR FILING DATE: 1997-06-06 |
| 40 | PRIOR APPLICATION NUMBER: 60/048,972 |
| 41 | PRIOR FILING DATE: 1997-06-06 |
| 42 | PRIOR APPLICATION NUMBER: 60/048,916 |
| 43 | PRIOR FILING DATE: 1997-06-06 |
| 44 | PRIOR APPLICATION NUMBER: 60/049,373 |
| 45 | PRIOR FILING DATE: 1997-06-06 |
| 46 | PRIOR APPLICATION NUMBER: 60/048,875 |
| 47 | PRIOR FILING DATE: 1997-06-06 |
| 48 | PRIOR APPLICATION NUMBER: 60/049,374 |
| 49 | PRIOR FILING DATE: 1997-06-06 |
| 50 | PRIOR APPLICATION NUMBER: 60/048,917 |
| 51 | PRIOR FILING DATE: 1997-06-06 |
| 52 | PRIOR APPLICATION NUMBER: 60/048,949 |
| 53 | PRIOR FILING DATE: 1997-06-06 |
| 54 | PRIOR APPLICATION NUMBER: 60/048,974 |
| 55 | PRIOR FILING DATE: 1997-06-06 |
| 56 | PRIOR APPLICATION NUMBER: 60/048,883 |
| 57 | PRIOR FILING DATE: 1997-06-06 |
| 58 | PRIOR APPLICATION NUMBER: 60/048,897 |
| 59 | PRIOR FILING DATE: 1997-06-06 |
| 60 | PRIOR APPLICATION NUMBER: 60/048,898 |
| 61 | PRIOR FILING DATE: 1997-06-06 |
| 62 | PRIOR APPLICATION NUMBER: 60/048,962 |
| 63 | PRIOR FILING DATE: 1997-06-06 |
| 64 | PRIOR APPLICATION NUMBER: 60/048,963 |
| 65 | PRIOR FILING DATE: 1997-06-06 |
| 66 | PRIOR APPLICATION NUMBER: 60/048,877 |
| 67 | PRIOR FILING DATE: 1997-06-06 |
| 68 | PRIOR APPLICATION NUMBER: 60/048,878 |
| 69 | PRIOR FILING DATE: 1997-06-06 |
| 70 | PRIOR APPLICATION NUMBER: 60/068,054 |
| 71 | PRIOR FILING DATE: 1997-12-18 |
| 72 | PRIOR APPLICATION NUMBER: 60/068,064 |
| 73 | PRIOR FILING DATE: 1997-12-18 |

; PRIOR APPLICATION NUMBER: 60/068,053
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/070,923
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/073,160
 ; PRIOR FILING DATE: 1998-01-30
 ; PRIOR APPLICATION NUMBER: 60/073,159
 ; PRIOR FILING DATE: 1998-01-30
 ; PRIOR APPLICATION NUMBER: 60/073,165
 ; PRIOR FILING DATE: 1998-01-30
 ; PRIOR APPLICATION NUMBER: 60/073,164
 ; PRIOR FILING DATE: 1998-01-30
 ; PRIOR APPLICATION NUMBER: 60/085,925
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,921
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,923
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,922
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/092,921
 ; PRIOR FILING DATE: 1998-07-15
 ; PRIOR APPLICATION NUMBER: 60/094,657
 ; PRIOR FILING DATE: 1998-07-30
 ; NUMBER OF SEQ ID NOS: 1245
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 97
 ; LENGTH: 1985
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (332)
 ; OTHER INFORMATION: n equals a.t.g, or c
 US-09-333-767-97

Query Match 23.2%; Score 30.6; DB 10; Length 1985;
 Best Local Similarity 52.8%; Pred. No. 0.46; Indels 0; Gaps 0;
 Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
 QY 4 CCTGTCTACTAGTCCCTTACCGAGTAAGGAAATGTACCGGACTTTTGGCGGAGAGGT 63
 Db 1523 CCTTCACTCCACGACCTTCCCAACAGGATAAGCTGGATCCCTTGGCCTTCTGAATAT 1582
 QY 64 GAAAGTCTAGTCCCTTCCGAACTCCAGTAAACGTGTGCTTCCCTAAGAGCGCAACAAT 123
 Db 1583 CCCAGTGTCTTTCAGGTTTCCCAAGACCACTTCCCTGTGGGCTTCCAAAATGGCCTTTATC 1642
 QY 124 ATTAC 128
 Db 1643 ATTTC 1647

RESULT 7
 US-10-004-860-97
 ; Sequence 97, Application US/10004860
 ; Publication No. US20030065160A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young et al.
 ; TITLE OF INVENTION: 207 Human Secreted Proteins
 ; FILE REFERENCE: P2007P1
 ; CURRENT APPLICATION NUMBER: US/10/004,860
 ; CURRENT FILING DATE: 2001-12-07
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 1227
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 97
 ; LENGTH: 1985
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE

; OTHER INFORMATION: n equals a.t.g, or c
 US-10-004-860-97
 Query Match 23.2%; Score 30.6; DB 13; Length 1985;
 Best Local Similarity 52.8%; Pred. No. 0.46; Indels 0; Gaps 0;
 Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
 QY 4 CCTGTCTACTAGTCCCTTACCGAGTAAGGAAATGTACCGGACTTTTGGCGGAGAGGT 63
 Db 1523 CCTTCACTCCACGACCTTCCCAACAGGATAAGCTGGATCCCTTGGCCTTCTGAATAT 1582
 QY 64 GAAAGTCTAGTCCCTTCCGAACTCCAGTAAACGTGTGCTTCCCTAAGAGCGCAACAAT 123
 Db 1583 CCCAGTGTCTTTCAGGTTTCCCAAGACCACTTCCCTGTGGGCTTCCAAAATGGCCTTTATC 1642
 QY 124 ATTAC 128
 Db 1643 ATTTC 1647

RESULT 8
 US-10-023-282-97
 ; Sequence 97, Application US/10023282
 ; Publication No. US20030092893A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young et al.
 ; TITLE OF INVENTION: 207 Human Secreted Proteins
 ; FILE REFERENCE: P2007P1
 ; CURRENT APPLICATION NUMBER: US/10/023,282
 ; CURRENT FILING DATE: 2001-12-20
 ; EARLIER APPLICATION NUMBER: 09/205,258
 ; EARLIER FILING DATE: 1998-12-04
 ; EARLIER APPLICATION NUMBER: PCT/US98/11422
 ; EARLIER FILING DATE: 1998-06-04
 ; EARLIER APPLICATION NUMBER: 60/048,885
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,375
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,881
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,880
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,896
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,020
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,876
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,895
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,884
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,894
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,971
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,964
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,882
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,899
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,893
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,900
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,901
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,992
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,915
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,019

APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA131P1
CURRENT APPLICATION NUMBER: US/10/264,237
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/16450
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: US 60/205,515
PRIOR FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 2876
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 115
LENGTH: 1987
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (5)-(5)
OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-115

Query Match 23.2%; Score 30.6; DB 16; Length 1987;
Best Local Similarity 52.8%; Pred. No. 0.46;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 4 CTTGTCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAAGGT 63
Db 1528 CTTCACTCCCGACCTTCCCAACAGGATAAGCTGGATCCCTTGGCCTTCTGAATAT 1587
Qy 64 GAAAGTCTCAGTCGCTTCCGAACTCCACGTAACGCTGCTCTTAAGACGGCAACAAT 123
Db 1588 CCAGTGTCTTCCAGTTTCCCAAGACCCTTCCCTGTGGGCTTCCAAATGGCCTTATC 1647
Qy 124 ATTAC 128
Db 1648 ATTTC 1652

RESULT 10
US-10-369-493-33481
Sequence 33481, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/350,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 33481
LENGTH: 2301
TYPE: DNA
ORGANISM: Desulfitobacterium hafniense
US-10-369-493-33481

Query Match 22.9%; Score 30.2; DB 16; Length 2301;
Best Local Similarity 56.6%; Pred. No. 0.69;
Matches 56; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 21 TTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAGGTGAAGTCTCAGTCGCTT 80
Db 2108 TCACTCTGAGAGGCAATGTTCTGCCCATTTGGCGAGTGAAGAAAAGTTCTCGCGCTC 2167
Qy 81 TCCGAAGTCCACGTAACGCTGCTCTTAAGACGGCAAA 119
Db 2168 ACCGGCGGGAATTAACAGGTTATTTTGCCTGAGCAAA 2206

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 97
LENGTH: 1985
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (332)
OTHER INFORMATION: n equals a,t,g, or c
US-10-023-282-97

Query Match 23.2%; Score 30.6; DB 15; Length 1985;
Best Local Similarity 52.8%; Pred. No. 0.46;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 4 CTTGTCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAAGGT 63
Db 1523 CTTCACTCCCGACCTTCCCAACAGGATAAGCTGGATCCCTTGGCCTTCTGAATAT 1582
Qy 64 GAAAGTCTCAGTCGCTTCCGAACTCCACGTAACGCTGCTCTTAAGACGGCAACAAT 123
Db 1583 CCAGTGTCTTCCAGGTTTCCCAAGACCCTTCCCTGTGGGCTTCCAAAATGGCCTTTATC 1642
Qy 124 ATTAC 128
Db 1643 ATTTC 1647

RESULT 9
US-10-264-237-115
Sequence 115, Application US/10264237
Publication No. US20040009491A1
GENERAL INFORMATION:

```
RESULT 11
US-09-887-576-863/c
; Sequence 863, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 863
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-576-863

Query Match      21.1%; Score 27.8; DB 9; Length 2000;
Best Local Similarity 55.8%; Pred. No. 5.6;
Matches 53; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 11 TACTAGTCCTTACCGAGTAAGAAATGTACCGGACTTTTGGCGGAGAAAGTGAAAGTC 70
Db 1780 TACAACCTCTCTTTTATACAAATACACTTCTTTCARAACTTGAAGCGGAATGTT 1721

Qy 71 TCACTGCTTTCCGAACCTCACGTAAACGTGCT 105
Db 1720 TTACGTGTTTGTGAACCTGACTGAAATGTTGAT 1686

RESULT 12
US-10-437-963-37892/c
; Sequence 37892, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 37892
; LENGTH: 4565
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_41579C.1
US-10-437-963-37892

Query Match      21.1%; Score 27.8; DB 17; Length 4565;
```

```
Best Local Similarity 62.0%; Pred. No. 7.5;
Matches 44; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 48 TTTCGGCGGAGAGTGAAAGTCTCAGTGCCTTTCCGAACCTCCACGTAACGTTGCTCTCC 107
Db 3628 TTTTCCTTGAGAGTTTGAAATGCTCTCTCACACTACGGACTCCACACAAATTTTGTTC 3569

Qy 108 TAAGACGGCAA 118
Db 3568 TTCTTCAGCAA 3558

RESULT 13
US-10-087-192-1786/c
; Sequence 1786, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 05/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1786
; LENGTH: 35548
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1786

Query Match      20.6%; Score 27.2; DB 13; Length 35548;
Best Local Similarity 53.8%; Pred. No. 26;
Matches 56; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 17 TCCTTACCGAGTAAGAAATGTACCGGACTTTTGGCGGAGAGGTGAAAGTCTCAGTC 76
Db 6033 TTCTCTACCGAGTTTCAGAGATGCCCGAGGTACTCGAGCATTCAGGGGAATGTTTCTC 5974

Qy 77 GCTTTCGAACTCCACGTAACGTGTGCTCCTTAAGACGGCAAC 120
Db 5973 TTTGGCCACAGGTGGCAGATGTGTGTCAGTGACAGCCAGCAGAC 5930

RESULT 14
US-10-027-632-38125/c
; Sequence 38125, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```

```
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38125
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-38125

Query Match          20.3%; Score 26.8; DB 13; Length 409;
Best Local Similarity 55.3%; Pred. No. 7.8; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 42;

QY      6 TTGCTTACTAGTCCCTTACCGAGTAAGGAAATGTACCGGACTTTTGGCGGAGAAAGGTGA 65
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      321 TTGCTTACTTAAGAATTCCAAAGTAATAAACTCACCAATGATTGGCTCTTTGAAATAAC 262
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      66 AGCTTCAGTCGCTTCCGAACTCCACGTAACG 99
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      261 AAGTCTTTGTCTTTTCATGAACCTCCGTTTAAATG 228
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 15
US-10-027-632-38125/c
; Sequence 38125, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/195,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38125
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-38125
```

```
Query Match          20.3%; Score 26.8; DB 16; Length 409;
Best Local Similarity 55.3%; Pred. No. 7.8; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 42;

QY      6 TTGCTTACTAGTCCCTTACCGAGTAAGGAAATGTACCGGACTTTTGGCGGAGAAAGGTGA 65
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      321 TTGCTTACTTAAGAATTCCAAAGTAATAAACTCACCAATGATTGGCTCTTTGAAATAAC 262
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      66 AGCTTCAGTCGCTTCCGAACTCCACGTAACG 99
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      261 AAGTCTTTGTCTTTTCATGAACCTCCGTTTAAATG 228
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: October 11, 2004, 12:14:32
Job time : 217 secs

OM nucleic - nucleic search, using sw model

Run on: October 11, 2004, 08:56:11 ; Search time 1398.5 Seconds
(without alignments)
2818.597 Million cell updates/sec

Title: US-10-009-317A-30
Perfect score: 132
Sequence: 1 ggaactgtctactagtcccc.....cggcaacaataattacagt 132

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vri:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| C 1 | 32.2 | 24.4 | 1201 | 13 BX338407 | BX338407 BX338407 |
| C 2 | 30.8 | 23.3 | 193 | 10 BF088574 | BF088574 RCI-HT088 |
| C 3 | 30.8 | 23.3 | 679 | 28 B2475082 | B2475082 B0NBK72TF |
| C 4 | 30.8 | 23.3 | 1007 | 29 C0NS015QE | AL105776 Drosophil |

| | | | | | | |
|------|------|------|-----|----|----------|--------------------|
| 5 | 30.6 | 23.2 | 320 | 12 | BI052058 | BI052058 RC3-GN026 |
| 6 | 30.6 | 23.2 | 348 | 12 | BI052068 | BI052068 RC3-GN026 |
| 7 | 30.6 | 23.2 | 365 | 12 | BI052060 | BI052060 RC3-GN026 |
| C 8 | 30.6 | 23.2 | 402 | 10 | AW937927 | AW937927 OVO-DT004 |
| C 9 | 30.6 | 23.2 | 431 | 9 | AI827281 | AI827281 wk90f12.x |
| C 10 | 30.6 | 23.2 | 451 | 9 | AA593405 | AA593405 nn37h03.s |
| C 11 | 30.6 | 23.2 | 453 | 10 | AW439761 | AW439761 hb77a08.x |
| C 12 | 30.6 | 23.2 | 460 | 10 | AW247815 | AW247815 2820449.3 |
| C 13 | 30.6 | 23.2 | 467 | 10 | AW129619 | AW129619 xe19t08.x |
| C 14 | 30.6 | 23.2 | 467 | 10 | AW150166 | AW150166 xg36c04.x |
| C 15 | 30.6 | 23.2 | 468 | 10 | BE220183 | BE220183 hv69b10.x |
| C 16 | 30.6 | 23.2 | 477 | 9 | AI393092 | AI393092 tg25h01.x |
| C 17 | 30.6 | 23.2 | 489 | 10 | AW474818 | AW474818 xy20f11.x |
| C 18 | 30.6 | 23.2 | 490 | 9 | AW071089 | AW071089 wv97b06.x |
| C 19 | 30.6 | 23.2 | 494 | 10 | BF197439 | BF197439 hr80f02.x |
| C 20 | 30.6 | 23.2 | 506 | 13 | BQ130065 | BQ130065 i74b01.x |
| C 21 | 30.6 | 23.2 | 509 | 9 | AI937084 | AI937084 wp72h04.x |
| C 22 | 30.6 | 23.2 | 510 | 12 | EM272572 | EM272572 ig99g05.x |
| C 23 | 30.6 | 23.2 | 519 | 12 | BI753090 | BI753090 g03025931 |
| C 24 | 30.6 | 23.2 | 542 | 14 | CA866048 | CA866048 i137d12.x |
| C 25 | 30.6 | 23.2 | 565 | 9 | AU148219 | AU148219 AU148219 |
| C 26 | 30.6 | 23.2 | 566 | 12 | BQ006052 | BQ006052 UI-H-E11 |
| C 27 | 30.6 | 23.2 | 567 | 9 | AU154596 | AU154596 AU154596 |
| C 28 | 30.6 | 23.2 | 570 | 9 | AA156048 | AA156048 z050c08.s |
| C 29 | 30.6 | 23.2 | 575 | 12 | BM671661 | BM671661 UI-E-CQ1- |
| C 30 | 30.6 | 23.2 | 605 | 12 | BG325061 | BG325061 602423503 |
| C 31 | 30.6 | 23.2 | 614 | 13 | BQ549701 | BQ549701 iK90809.x |
| C 32 | 30.6 | 23.2 | 619 | 10 | BE876480 | BE876480 601486221 |
| C 33 | 30.6 | 23.2 | 619 | 12 | BI091965 | BI091965 602856754 |
| C 34 | 30.6 | 23.2 | 635 | 14 | CA423376 | CA423376 UI-H-FL0- |
| C 35 | 30.6 | 23.2 | 639 | 10 | AW473070 | AW473070 xx79f03.x |
| C 36 | 30.6 | 23.2 | 714 | 12 | BM996408 | BM996408 UI-H-DT0- |
| C 37 | 30.6 | 23.2 | 731 | 9 | AI922738 | AI922738 w013b05.x |
| C 38 | 30.6 | 23.2 | 738 | 13 | BU618317 | BU618317 UI-H-FH1- |
| C 39 | 30.6 | 23.2 | 742 | 14 | CB250483 | CB250483 UI-CF-FN0 |
| C 40 | 30.6 | 23.2 | 751 | 13 | BQ576177 | BQ576177 UI-H-EZ1- |
| C 41 | 30.6 | 23.2 | 759 | 13 | BU633104 | BU633104 UI-H-FL1- |
| C 42 | 30.6 | 23.2 | 791 | 14 | CD367608 | CD367608 UI-H-FT1- |
| C 43 | 30.6 | 23.2 | 826 | 12 | BG680371 | BG680371 602629268 |
| C 44 | 30.6 | 23.2 | 888 | 9 | AL522827 | AL522827 AL522827 |
| C 45 | 30.6 | 23.2 | 902 | 13 | EX362182 | EX362182 BX362182 |

ALIGNMENTS

RESULT 1
BX338407/c
LOCUS BX338407 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI059YK02 3-PRIME, mRNA sequence.
ACCESSION BX338407.1 GI:30335707
VERSION BX338407
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li W.B., Gruber C., Jessee J., and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3398.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DI059BF01NPI&cluster=3398.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600

```

FEATURES
  source
    Faraday Avenue Genoscope sequence ID : CS0D1059BF01NP1.
    Location/Qualifiers
      1..1201
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0D1059YK02"
        /tissue="Placenta"
        /clone_lib="Homo sapiens Placenta COT 25-NORMALIZED"
        /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
  Query Match      24.4%; Score 32.2; DB 13; Length 1201;
  Best Local Similarity 53.6%; Pred. No. 5.3;
  Matches 57; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 4 CTTGTCTACTAGTCCTTACCGAGTAAGAAATGTACCGAGCTTTTGGCGGAGAGGT 63
Db 401 CTTGACTCCAGCACCTTGCCTCCCAACAGGATAGCTGGATCCCTTTGGGCTTCTGAATAT 342
QY 64 GAAGTCTAGTGCCTTCCGACTCCACGTACACGTGCTCTTAAGACGGCAACAAT 123
Db 341 CCAGTGTCTTTCAGTTTCCCAAGACCACCTTCCCTGTGGCTTCCAAATGGCCATTATC 282
QY 124 ATTAC 128
Db 281 ATTTC 277

RESULT 2
BF088574/c
LOCUS
  DEFINITION R01-HT0881-130900-014-b11 HT0881 Homo sapiens cDNA, mRNA sequence.
  ACCESSION BF088574
  VERSION BF088574.1 GI:10894284
  KEYWORDS EST.
  SOURCE Homo sapiens (human)
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 193)
    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
    Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
    Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
    Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
    O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
    Simpson,A.J.
  Shotgun sequencing of the human transcriptome with ORF expressed
  sequence tags
  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  20202663
  10737800
  COMMENT
    Contact: Simpson A.J.G.
    Laboratory of Cancer Genetics
    Ludwig Institute for Cancer Research
    Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
    Brazil
    Tel: +55-11-2704922
    Fax: +55-11-2707001
    Email: asimpson@ludwig.org.br
    This sequence was derived from the FAPESP/LICR Human Cancer Genome
    Project. This entry can be seen in the following URL
    (http://www.ludwig.org.br/scripts/gethtml2.pl?title=et2=R01-HT0881-130
    900-014-b11&t3=2000-09-13&t4=1)
    Seq primer: puc 18 forward
    High quality sequence start: 28
    High quality sequence stop: 192.
    Location/Qualifiers
      1..193
        /organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0881"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
  Query Match      23.3%; Score 30.8; DB 10; Length 193;
  Best Local Similarity 55.7%; Pred. No. 8.3;
  Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 3 ACCTTGCTCTACTAGTCCTTACCGAGTAAGAAATGTACCGAGCTTTTGGCGGAGAGG 62
Db 111 ACCTCAACTCTGCTCAATCTGATGGAAGTAATCTGCTTAACAATCTCAGAAGGACTG 52
QY 63 TGAAGTCTCAGTCGCTTTCCGAACCTCCACGTAAACCTGTGCTCTCT 108
Db 51 TGCAAGTCATGTCGCTGTGTGATTCATCTCATGAAACGAGGATCTCT 6

RESULT 3
BZ475082
LOCUS
  DEFINITION B01.6.2 KB tot Brassica oleracea genomic clone B01.6.2,
  genomic survey sequence.
  ACCESSION BZ475082
  VERSION BZ475082.1 GI:26775627
  KEYWORDS GSS.
  SOURCE Brassica oleracea
  ORGANISM Brassica oleracea
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
  1 (bases 1 to 679)
    Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
    Whole genome shotgun sequencing of Brassica oleracea
    Unpublished (2001)
    Other GSSs: B01.6.2 KB tot Brassica oleracea
    Contact: Chris Town
    TIGR
    9712 Medical Center Drive, Rockville, MD 20850, USA.
    Tel: 301-838-3523
    Fax: 301-838-0208
    Email: cdtown@tigr.org
    DNA is from a doubled haploid provided by Tom Osborn.
    Seq primer: TF
    Class: sheared ends.
    Location/Qualifiers
      1..679
        /organism="Brassica oleracea"
        /mol_type="genomic DNA"
        /strain="TO1000DH3"
        /db_xref="taxon:3712"
        /clone="B01.6.2 KB tot"
        /clone_lib="BO 1.6.2 KB tot"
        /note="Vector: pUC18; Site 1: BstXI; 1.6-2 kb sheared
        total DNA inserted into pUC18 using BstXI linkers"

FEATURES
  source
    Query Match      23.3%; Score 30.8; DB 28; Length 679;
    Best Local Similarity 55.7%; Pred. No. 13;
    Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 17 TCCCTTACCGAGTAAGAAATGTACCGAGCTTTTGGCGGAGAGGTCGAAGTCTCAGTC 76
Db 15 TCCCACTAATTTTCAGAGACACAAATCGGCTTTTGTGCGAGAGGTAATTTTCGGATA 74

```


O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC3&t2=RC3-GN0268-170101-012-h08&t3=2001-01-17&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 6
High quality sequence stop: 320.
Location/Qualifiers
1. 320
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_xref="Adult"
/clone_lib="GN0268"
/notes="Organ: placenta normal; Vector: puc18; Site: 1; Smal; Site 2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 23.2%; Score 30.6; DB 12; Length 320;
Best Local Similarity 52.8%; Pred. No. 12;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

4 CTTTGCTACTAGTCCCTTACGAGTAAGGAAATGTCGGGACTTTTCGGCGGAGAAAGT 63
161 CTTTCACTCCAGACGCTTCCGCAACAGAGATAGCTGCCCTTGGCCTTCTGAATAT 220
64 GAAAGTCTCAGTCGCTTCCGAACTCCAGTAAAGCTGTGCTCTTAAGACGGCAACAAT 123
221 CCAGTGTCTTCAGGTTTCCAGACCACCTCCCTGTGGGCTTCNAAATGCGCTTATC 280
124 ATTAC 128
281 ATTTC 285

RESULT 6
BI052068
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 349)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., deOliveira,A.F., Matsuikuma,A., Bata,G.S., Simpson,D.H., Brundstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT

CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=RC3&t2=RC3-GN0268-170101-012-a07&t3=2001-01-17&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 33
High quality sequence stop: 348.

FEATURES
source
1..348
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0268"
/notes="Organ: placenta normal; Vector: puc18; Site 1: Sma1; Site 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 23.2%; Score 30.6; DB 12; Length 348;
Best Local Similarity 52.8%; Pred. No. 12;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CCTGTCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAGGT 63
|||
Db 189 CCTTCACTCCAGACCTTCCCAACAGGATAGCTGGATCCCTTGGCTTCTGAATAT 248
|||
QY 64 GAAAGTCTAGTCTTCGAACTCCACGTAAACGTGTCTCTTAAGACGGCAACAAT 123
|||
Db 249 CCCAGTGTCTTCAGGTTTCCCAAGACCACTTCCCTGTGGGCTTCCAAATGGCCTTATC 308
|||
QY 124 ATTAC 128
|||
Db 309 ATTTC 313
|||

RESULT 7
BI052060 365 bp mRNA linear EST 15-JUN-2001
LOCUS RC3-GN0268-170101-012-a07 GN0268 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BI052060
VERSION BI052060.1 GI:14459590
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 365)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT

CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=RC3&t2=RC3-GN0268-170101-012-a07&t3=2001-01-17&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 51
High quality sequence stop: 365.

FEATURES
source
1..365
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0268"
/notes="Organ: placenta normal; Vector: puc18; Site 1: Sma1; Site 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 23.2%; Score 30.6; DB 12; Length 365;
Best Local Similarity 52.8%; Pred. No. 12;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CCTGTCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAGGT 63
|||
Db 206 CCTTCACTCCAGACCTTCCCAACAGGATAGCTGGATCCCTTGGCTTCTGAATAT 265
|||
QY 64 GAAAGTCTCAGTCGCTTCCGAACTCCACGTAAACGTGTCTCTTAAGACGGCAACAAT 123
|||
Db 266 CCCAGTGTCTTCAGGTTTCCCAAGACCACTTCCCTGTGGGCTTCCAAATGGCCTTATC 325
|||
QY 124 ATTAC 128
|||
Db 326 ATTTC 330
|||

RESULT 8
AW937927/c 402 bp mRNA linear EST 30-MAY-2000
LOCUS QV0-DT0047-170200-122-h12 DT0047 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION AW937927
VERSION AW937927.1 GI:8113357
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 402)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

10737800
 PUBMED
 COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=st2-QV0-DT0047-170
 200-122-h12&t=2000-02-17&t4=1)

Seq primer: puc 18 forward
 High quality sequence start: 20
 High quality sequence stop: 402.

Location/Qualifiers
 1. 402
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="DT0047"
 /note="Organ: denis drash; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 23.2%; Score 30.6; DB 10; Length 402;
 Best Local Similarity 52.8%; Pred. No. 13;
 Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CTTGTCTACTAGTCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAAGGT 63
 |||||
 Db 168 CTTCACTCCGAGCCTTCCACAGGATAGCTGGATCCCTTGGCCTTCTGAATAT 109
 |||||

QY 64 GAAAGTCTCAGTCGTTTCCGACTCCAGTAAACGTGTGCTCTTAAGACGCAACAAT 123
 |||||
 Db 108 CCGAGTCTTCAGGTTTCCCAAGACCACTTCCCTGTGGGCTTCCAAAATGSCCTTTATC 49
 |||||

QY 124 ATTAC 128
 |||||
 Db 48 ATTTC 44

RESULT 9
 AI827281/c
 LOCUS
 DEFINITION wk90f12.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2422703 3',
 mRNA sequence.
 ACCESSION AI827281
 VERSION AI827281.1 GI:5447952
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 431)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 916 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 413.
 Location/Qualifiers
 1. 431
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="IMAGE:2422703"
 /tissue_type="squamous cell carcinoma, poorly
 differentiated (4 pooled tumors, including primary and
 metastatic)"
 /dev_stage="adult"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI_CGAP Lu19"
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 pooled lung tumor tissue, and was then primed with a Not I
 - oligo(dT) primer. Double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT7T3 vector. Library went through one round of
 normalization. Library constructed by Bento Soares and M.
 Fatima Bonaldo."

ORIGIN

Query Match 23.2%; Score 30.6; DB 9; Length 431;
 Best Local Similarity 52.8%; Pred. No. 13;
 Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CTTGTCTACTAGTCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAAGGT 63
 |||||
 Db 431 CTTCACTCCGAGCCTTCCCAAGACCACTTCCCTGTGGGCTTCCAAAATGSCCTTTATC 372
 |||||

QY 64 GAAAGTCTCAGTCGTTTCCGACTCCAGTAAACGTGTGCTCTTAAGACGCAACAAT 123
 |||||
 Db 371 CCGAGTCTTCAGGTTTCCAGACCACTTCCCTGTGGGCTTCCAAAATGSCCTTTATC 312
 |||||

QY 124 ATTAC 128
 |||||
 Db 311 ATTTC 307

RESULT 10
 AA593405/c
 LOCUS
 DEFINITION m57h03.81 NCI_CGAP_Kid6 Homo sapiens cDNA clone IMAGE:1088021 3',
 mRNA sequence.
 ACCESSION AA593405
 VERSION AA593405.1 GI:2409167
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 451)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Stratagene, Inc.
 DNA Sequencing by: Greg Lennon, Ph.D.
 cDNA Library Arrayed by: Washington University Genome Sequencing Center
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

10737800
 PUBMED
 COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=st2-QV0-DT0047-170
 200-122-h12&t=2000-02-17&t4=1)

Seq primer: puc 18 forward
 High quality sequence start: 20
 High quality sequence stop: 402.

Location/Qualifiers
 1. 402
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="DT0047"
 /note="Organ: denis drash; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 23.2%; Score 30.6; DB 10; Length 402;
 Best Local Similarity 52.8%; Pred. No. 13;
 Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CTTGTCTACTAGTCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAAGGT 63
 |||||
 Db 168 CTTCACTCCGAGCCTTCCACAGGATAGCTGGATCCCTTGGCCTTCTGAATAT 109
 |||||

QY 64 GAAAGTCTCAGTCGTTTCCGACTCCAGTAAACGTGTGCTCTTAAGACGCAACAAT 123
 |||||
 Db 108 CCGAGTCTTCAGGTTTCCCAAGACCACTTCCCTGTGGGCTTCCAAAATGSCCTTTATC 49
 |||||

QY 124 ATTAC 128
 |||||
 Db 48 ATTTC 44

RESULT 9
 AI827281/c
 LOCUS
 DEFINITION wk90f12.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2422703 3',
 mRNA sequence.
 ACCESSION AI827281
 VERSION AI827281.1 GI:5447952
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 431)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 916 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 413.
 Location/Qualifiers
 1. 431
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="IMAGE:2422703"
 /tissue_type="squamous cell carcinoma, poorly
 differentiated (4 pooled tumors, including primary and
 metastatic)"
 /dev_stage="adult"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI_CGAP Lu19"
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 pooled lung tumor tissue, and was then primed with a Not I
 - oligo(dT) primer. Double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT7T3 vector. Library went through one round of
 normalization. Library constructed by Bento Soares and M.
 Fatima Bonaldo."

ORIGIN

Query Match 23.2%; Score 30.6; DB 9; Length 431;
 Best Local Similarity 52.8%; Pred. No. 13;
 Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CTTGTCTACTAGTCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAAGGT 63
 |||||
 Db 431 CTTCACTCCGAGCCTTCCCAAGACCACTTCCCTGTGGGCTTCCAAAATGSCCTTTATC 372
 |||||

QY 64 GAAAGTCTCAGTCGTTTCCGACTCCAGTAAACGTGTGCTCTTAAGACGCAACAAT 123
 |||||
 Db 371 CCGAGTCTTCAGGTTTCCAGACCACTTCCCTGTGGGCTTCCAAAATGSCCTTTATC 312
 |||||

QY 124 ATTAC 128
 |||||
 Db 311 ATTTC 307

RESULT 10
 AA593405/c
 LOCUS
 DEFINITION m57h03.81 NCI_CGAP_Kid6 Homo sapiens cDNA clone IMAGE:1088021 3',
 mRNA sequence.
 ACCESSION AA593405
 VERSION AA593405.1 GI:2409167
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 451)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Stratagene, Inc.
 DNA Sequencing by: Greg Lennon, Ph.D.
 cDNA Library Arrayed by: Washington University Genome Sequencing Center
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

```
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 405.
Location/Qualifiers
1. .451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1088021"
/sex="mixed"
/tissue_type="kidney tumor"
/lab_host="NCI CGAP Kid6"
/notes="Organ: kidney; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'
GAATCGGCACGAG 3' 3' adaptor sequence: 5'
CTCAGATTTTCTTTTCTTTT 3' Average insert size: 1.0 kb."

ORIGIN
Query Match 23.2%; Score 30.6; DB 9; Length 451;
Best Local Similarity 52.8%; Pred. No. 13;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CCTGTCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAGGT 63
DB 399 CCTTCACTCCAGCACCTTGCCTCCACAGGATAGCTGGATCCCTTGGCCTTCTGAATAT 340
QY 64 GAAAGTCTCAGTCGCTTCCGAACTCCAGTAAACGTCCTCTTAAGACGCGCAACAAT 123
DB 339 CCAGTCTCTCAGTTCCAGACCACTTCCCTGTGGGCTTCCAAATGGCCTTTATC 280
QY 124 ATTAC 128
DB 279 ATTTC 275

RESULT 11
AW439761/c 453 bp mRNA linear EST 14-FEB-2000
LOCUS h577a08.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2889206 3',
DEFINITION mRNA sequence.
ACCESSION AW439761
VERSION AW439761.1 GI:6975067
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 453)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA sequencing by: Greg Lennon, Ph.D.
Cloning strategy: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 418.
Location/Qualifiers
1. .453
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2889206"
/tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"

FEATURES
source
1. .451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2889206"
/tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"

Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 405.
Location/Qualifiers
1. .451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1088021"
/sex="mixed"
/tissue_type="kidney tumor"
/lab_host="NCI CGAP Kid6"
/notes="Organ: kidney; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'
GAATCGGCACGAG 3' 3' adaptor sequence: 5'
CTCAGATTTTCTTTTCTTTT 3' Average insert size: 1.0 kb."

ORIGIN
Query Match 23.2%; Score 30.6; DB 9; Length 451;
Best Local Similarity 52.8%; Pred. No. 13;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CCTGTCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAGGT 63
DB 399 CCTTCACTCCAGCACCTTGCCTCCACAGGATAGCTGGATCCCTTGGCCTTCTGAATAT 340
QY 64 GAAAGTCTCAGTCGCTTCCGAACTCCAGTAAACGTCCTCTTAAGACGCGCAACAAT 123
DB 339 CCAGTCTCTCAGTTCCAGACCACTTCCCTGTGGGCTTCCAAATGGCCTTTATC 280
QY 124 ATTAC 128
DB 279 ATTTC 275

RESULT 11
AW439761/c 453 bp mRNA linear EST 14-FEB-2000
LOCUS h577a08.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2889206 3',
DEFINITION mRNA sequence.
ACCESSION AW439761
VERSION AW439761.1 GI:6975067
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 453)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA sequencing by: Greg Lennon, Ph.D.
Cloning strategy: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 418.
Location/Qualifiers
1. .453
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2889206"
/tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"

FEATURES
source
1. .451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2889206"
/tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"

Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 405.
Location/Qualifiers
1. .451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1088021"
/sex="mixed"
/tissue_type="kidney tumor"
/lab_host="NCI CGAP Kid6"
/notes="Organ: kidney; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'
GAATCGGCACGAG 3' 3' adaptor sequence: 5'
CTCAGATTTTCTTTTCTTTT 3' Average insert size: 1.0 kb."

ORIGIN
Query Match 23.2%; Score 30.6; DB 9; Length 451;
Best Local Similarity 52.8%; Pred. No. 13;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CCTGTCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAGGT 63
DB 399 CCTTCACTCCAGCACCTTGCCTCCACAGGATAGCTGGATCCCTTGGCCTTCTGAATAT 340
QY 64 GAAAGTCTCAGTCGCTTCCGAACTCCAGTAAACGTCCTCTTAAGACGCGCAACAAT 123
DB 339 CCAGTCTCTCAGTTCCAGACCACTTCCCTGTGGGCTTCCAAATGGCCTTTATC 280
QY 124 ATTAC 128
DB 279 ATTTC 275

RESULT 11
AW439761/c 453 bp mRNA linear EST 14-FEB-2000
LOCUS h577a08.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2889206 3',
DEFINITION mRNA sequence.
ACCESSION AW439761
VERSION AW439761.1 GI:6975068
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 460)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other ESTs: 2820449.5prime
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
Project Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center
Trimming: cross match from University of Washington Genome Center
PHRAP suite Poly-T Identification: patMatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu polyadenylation: Based upon the
presence of a XhoI site followed by a run of 14 or more T residues
at the beginning of the sequence, this cDNA insert was
polyadenylated.
Plate: L1CM4 row: D column: 18
High quality sequence stop: 326.
Location/Qualifiers
1. .460
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2820449"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_7"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
```

| | | | |
|----|-----|--|-----|
| Qy | 4 | CCTTGTCTACTAGTCCTCTTACCGAGTAAGGAANTGTACCGGACTTTTTCGGCGAGGAAGGT | 63 |
| Db | 435 | CCTTCACTCCCAACACCTTTGCCAACAGGATAGCTGGATCCCCCTTGGCCCTTTTGAATAT | 376 |
| Qy | 64 | GAAGAATCTCAGTCGCCTTTCGGAATCCAGTAAACGTGTGCTCCTAAGACGGCAAAACAAT | 123 |
| Db | 375 | CCCAAGTGTCTTCAGGTTTCCCAAGACACCTTCCCTGTGGCTTCCAAAATGGCCCTTATC | 316 |

| Qy | 124 | ATTAC | 128 |
|----|-----|-------|-----|
| Db | 315 | ATTTC | 311 |

RESULT 14
AW150166/C

| | | | | | |
|------------|--|--------|------|--------|-----------------|
| LOCUS | AW150166 | 467 bp | mRNA | linear | EST 03-NOV-1999 |
| DEFINITION | xg36c04.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:3629638 3', mRNA sequence. | | | | |
| ACCESSION | AW150166 | | | | |

accession AM150166.1
 version
 keywords
 source Homo sapiens (human)
 organism Homo sapiens

RESULT 13
AW129619/C

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
 Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

Samuel-Buck, M.D., F.R.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found at: <http://www.ncbi.nlm.nih.gov/CGAP/>

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -400P from Gibco
High quality sequence stock: 416

```

FEATURES
  source
    Location/Qualifiers
      1..467
      /organism="Homo sapiens"
  seq primer: -400P from Gbco
  High quality sequence stop: 416.

```

```
source
1. .467
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/ol_nucleotides=200000
```

```

/db_xref="taxon:9606"
/clone="IMAGE:2629638"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"

```

```
adenocarcinoma, 7 pooled tumors"
/lab host="DH10B"
/clone_lib="NCI CGAP Ut1"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; SalI;
```

/notes"Organ: uterus"; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .467 |

| | | | | |
|------------|-------------|-------------|--------|-------------|
| 11538-014" | 23.2% | Score 30.6; | DB 10; | Length 467; |
| ORIGIN | Query Match | | | |

Query Match 23.2%; Score 30.6; DB 10; Length 467;
Best Local Similarity 52.8%; Pred. No. 13;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Matches 66; conservative 0; mismatches 55; indels 0; gaps 0

Qy 4 CTTGTCTACTCTCTTACCAGTAAGAAATGACCGACTTTTTCGGCGAGAAAGGT 63

Dh 437 CTTTCACTCCACGACTTTGCCCAACAGGATAGCTGGATCCCTCTGGCCTTCTGAATAT 378

| | | | |
|----|-----|---|-----|
| Db | 437 | CC TTC ACTCCAGCACTTGGCCCAACAGGATAAGCTGGATCCCTTGGCCTTCTGAATAT | 378 |
| Qy | 64 | GA AAGTCTCAGTCGCTTTTCGAACTCCACGTAAACCTGTGCTCCTTAAGACGGCAACAAT | 123 |

ORIGIN
44372 040

| Qy | 124 | ATTAC | 128 |
|----|-----|-------|-----|
| Db | 317 | ATTTC | 313 |

RESULT 15
 BE220183/c
 LOCUS
 DEFINITION hv69b10.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3178651 3',
 mRNA sequence.
 ACCESSION BE220183
 VERSION BE220183
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 468)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL, send email to:
info@image.lnl.gov
 Seq primer: -400P from Gibco
 High quality sequence stop: 468.

FEATURES
 source
 1..468
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3178651"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Lu24"
 /notes="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; Plasmid DNA from the normalized
 library NCI_CGAP_Lu5 was prepared, and ss circles were
 made in vitro. Following HAP purification, this DNA was
 used as tracer in a subtractive hybridization reaction.
 The driver was PCR-amplified cDNAs from a pool of 5,000
 clones made from the same library (cloneIDs
 1414920-1417991 and 1520904-152439). Subtraction by Bento
 Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 23.2%; Score 30.6; DB 10; Length 468;
 Best Local Similarity 52.8%; Pred. No. 13;
 Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 4 CCTTGTTACTAGTCCCTTACCGAGTAAGCAATGTACCGGACTTTTGGCGGAGAGGT 63
 |||||
 Db 426 CCTTCACTCCAGCACCTTCCCAACAGGATAGCTGGATCCCTTTGGCCTTCTGAATAT 367
 |||||
 Qy 64 GAAAGTCTCAGTGGCTTCCGAACTCCACGCTAAACGTGTGCTCTTAAGACGGCAACAAT 123
 |||||
 Db 366 CCAGTGTCTTCAGGTTTCCCAAGACCACTTCCCTGTGGGCTTCCAAAATGGCCTTTATC 307
 |||||
 Qy 124 ATTAC 128
 |||||
 Db 306 ATTTC 302

Search completed: October 11, 2004, 10:55:04
 Job time : 1403.5 secs

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

OM nucleic - nucleic search, using sw model

Run on: October 11, 2004, 08:25:36 ; Search time 1124.5 Seconds
(without alignments)
5087.844 Million cell updates/sec

Title: US-10-009-317A-31

Perfect score: 132

Sequence: 1 ggacatctctactagtagtcccc.....cggttgacaattattccogca 132

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.ov.*
- 5: gb.pat.*
- 6: gb.ph.*
- 7: gb.pl.*
- 8: gb.pr.*
- 9: gb.ro.*
- 10: gb.sts.*
- 11: gb.sy.*
- 12: gb.un.*
- 13: gb.vi.*
- 14: gb.ba.*
- 15: em.fun.*
- 16: em.hum.*
- 17: em.in.*
- 18: em.mu.*
- 19: em.or.*
- 20: em.ov.*
- 21: em.pat.*
- 22: em.ph.*
- 23: em.pl.*
- 24: em.ro.*
- 25: em.sts.*
- 26: em.un.*
- 27: em.vi.*
- 28: em.htg.hum.*
- 29: em.htg.inv.*
- 30: em.htg.other.*
- 31: em.htg.mus.*
- 32: em.htg.pln.*
- 33: em.htg.fed.*
- 34: em.htg.mam.*
- 35: em.htg.vrt.*
- 36: em.sy.*
- 37: em.htgo.hum.*
- 38: em.htgo.mus.*
- 39: em.htgo.pln.*
- 40: em.htgo.fed.*
- 41: em.htgo.mam.*
- 42: em.htgo.vrt.*

Pred. No. is the number of results predicted by chance to have a

| Result No. | Score | Match % | Length | DB | ID | Description |
|------------|-------|---------|--------|----|-----------|--------------------|
| 1 | 132 | 100.0 | 132 | 6 | ARI181805 | ARI181805 Sequence |
| 2 | 132 | 100.0 | 132 | 6 | AX047117 | AX047117 Sequence |
| 3 | 109.2 | 82.7 | 132 | 6 | ARI181804 | ARI181804 Sequence |
| 4 | 109.2 | 82.7 | 132 | 6 | AX047116 | AX047116 Sequence |
| 5 | 32.6 | 24.7 | 178965 | 10 | AL844144 | AL844144 Mouse DNA |
| 6 | 32.6 | 24.7 | 241189 | 2 | AC107644 | AC107644 Mus muscu |
| 7 | 30.4 | 23.0 | 5339 | 10 | AK129480 | AK129480 Mus muscu |
| 8 | 30.4 | 23.0 | 207736 | 2 | AC120854 | AC120854 Rattus no |
| 9 | 30.4 | 23.0 | 216422 | 2 | AC123263 | AC123263 Rattus no |
| 10 | 30.4 | 23.0 | 239297 | 10 | AC119848 | AC119848 Mus muscu |
| 11 | 30.2 | 22.9 | 176183 | 10 | AL589766 | AL589766 Mouse DNA |
| 12 | 30.2 | 22.9 | 191884 | 10 | BX248984 | BX248984 Mouse DNA |
| 13 | 30.2 | 22.9 | 194807 | 10 | AC129542 | AC129542 Mus muscu |
| 14 | 30 | 22.7 | 193304 | 2 | AC118868 | AC118868 Rattus no |
| 15 | 30 | 22.7 | 239140 | 2 | AC099298 | AC099298 Rattus no |
| 16 | 29.8 | 22.6 | 1646 | 3 | AY331805 | AY331805 Ephelota |
| 17 | 29.8 | 22.6 | 39910 | 2 | AC104217 | AC104217 Mus muscu |
| 18 | 29.8 | 22.6 | 100000 | 9 | AP000066 | AP000066 Homo sapi |
| 19 | 29.8 | 22.6 | 104996 | 9 | AP006291 | AP006291 Homo sapi |
| 20 | 29.8 | 22.6 | 153973 | 2 | AC138373 | AC138373 Homo sapi |
| 21 | 29.8 | 22.6 | 176794 | 9 | AC137579 | AC137579 Homo sapi |
| 22 | 29.8 | 22.6 | 179829 | 2 | AC137690 | AC137690 Homo sapi |
| 23 | 29.8 | 22.6 | 190274 | 9 | AC138356 | AC138356 Homo sapi |
| 24 | 29.8 | 22.6 | 216415 | 10 | AL691423 | AL691423 Mouse DNA |
| 25 | 29.8 | 22.6 | 300620 | 1 | AE016782 | AE016782 Pseudomon |
| 26 | 29.6 | 22.4 | 3112 | 3 | AY113357 | AY113357 Drosophil |
| 27 | 29.6 | 22.4 | 43341 | 2 | AC014471 | AC014471 Drosophil |
| 28 | 29.6 | 22.4 | 151610 | 3 | AC009462 | AC009462 Drosophil |
| 29 | 29.6 | 22.4 | 155840 | 2 | AC006495 | AC006495 Drosophil |
| 30 | 29.6 | 22.4 | 199044 | 3 | AC009394 | AC009394 Drosophil |
| 31 | 29.6 | 22.4 | 224890 | 3 | AE003721 | AE003721 Drosophil |
| 32 | 29.4 | 22.3 | 80019 | 8 | ATF25E4 | AL050399 Arabidops |
| 33 | 29.4 | 22.3 | 198788 | 8 | ATCHRIV32 | AL164532 Arabidops |
| 34 | 29.2 | 22.1 | 14179 | 1 | AE007053 | AE007053 Mycobacte |
| 35 | 29.2 | 22.1 | 82993 | 6 | AX704274 | AX704274 Sequence |
| 36 | 29.2 | 22.1 | 130981 | 2 | AC087702 | AC087702 Trypanoso |
| 37 | 29.2 | 22.1 | 180393 | 9 | AC015938 | AC015938 Homo sapi |
| 38 | 29.2 | 22.1 | 181991 | 9 | AC025882 | AC025882 Homo sapi |
| 39 | 29.2 | 22.1 | 200933 | 10 | AL805910 | AL805910 Mouse DNA |
| 40 | 29.2 | 22.1 | 291050 | 1 | EX248340 | EX248340 Mycobacte |
| 41 | 29.2 | 22.1 | 308354 | 2 | AC093482 | AC093482 Mus muscu |
| 42 | 29.2 | 22.1 | 346186 | 1 | EX842578 | EX842578 Mycobacte |
| 43 | 29 | 22.0 | 2004 | 6 | BD179694 | BD179694 Highly th |
| 44 | 29 | 22.0 | 112732 | 2 | AC092367 | AC092367 Homo sapi |
| 45 | 29 | 22.0 | 166098 | 9 | AC020641 | AC020641 Homo sapi |

ALIGNMENTS

RESULT 1

ARI181805

LOCUS ARI181805

DEFINITION Sequence 31 from patent US 6335318.

ACCESSION ARI181805

VERSION ARI181805.1

KEYWORDS GI:20224019

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 132)

AUTHORS Selsted,M.E., Tang,Y.-Q., Yuan,J. and Ouellette,A.J.

TITLE Antimicrobial theta defensins and methods of using same

JOURNAL Patent: US 6335318-A 31 01-JAN-2002;

FEATURES Location/Qualifiers

linear PAT 20-APR-2002

source 1. .132
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 4.4e-34;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTGCCGACTTTTGGCGGAGAA 60
DB 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTGCCGACTTTTGGCGGAGAA 60

QY 61 GGCGAAAGTCTCAGTCCGCTCTCCGAACTCCACGAAACGCTGCTCTCTCAACGGTTGAC 120
DB 61 GGCGAAAGTCTCAGTCCGCTCTCCGAACTCCACGAAACGCTGCTCTCTCAACGGTTGAC 120

QY 121 AATATTCCGCA 132
DB 121 AATATTCCGCA 132

RESULT 2
AX047117
LOCUS AX047117 132 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 31 from Patent WO0068265.
ACCESSION AX047117
VERSION AX047117.1 GI:11876503
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1
AUTHORS Selsted,M.E., Tang,Y.Q., Yuan,J. and Ouellette,A.J.
TITLE Antimicrobial theta defensins and methods of using same
JOURNAL Patent: WO 0068265-A 31 16-NOV-2000;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA (US)
FEATURES Location/Qualifiers
source 1. .132
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic Construct"

ORIGIN

Query Match 100.0%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 4.4e-34;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTGCCGACTTTTGGCGGAGAA 60
DB 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTGCCGACTTTTGGCGGAGAA 60

QY 61 GGCGAAAGTCTCAGTCCGCTCTCCGAACTCCACGAAACGCTGCTCTCTCAACGGTTGAC 120
DB 61 GGCGAAAGTCTCAGTCCGCTCTCCGAACTCCACGAAACGCTGCTCTCTCAACGGTTGAC 120

QY 121 AATATTCCGCA 132
DB 121 AATATTCCGCA 132

RESULT 3
AR181804
LOCUS AR181804 132 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 30 from patent US 633518.
ACCESSION AR181804
VERSION AR181804.1 GI:20224018
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 132)

AUTHORS Selsted,M.E., Tang,Y.Q., Yuan,J. and Ouellette,A.J.
TITLE Antimicrobial theta defensins and methods of using same
JOURNAL Patent: US 633518-A 30 01-JAN-2002;
FEATURES Location/Qualifiers
source 1. .132
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 82.7%; Score 109.2; DB 6; Length 132;
Best Local Similarity 90.0%; Pred. No. 2.7e-26;
Matches 117; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTGCCGACTTTTGGCGGAGAA 60
DB 1 GGACCTTGTCTACTAGTCCCTTACCGAGTAAGAAATGTACCAGACTTTTGGCGGAGAA 60

QY 61 GGCGAAAGTCTCAGTCCGCTCTCCGAACTCCACGAAACGCTGCTCTCTCAACGGTTGAC 120
DB 61 GGCGAAAGTCTCAGTCCGCTCTCCGAACTCCACGAAACGCTGCTCTCTCAACGGTTGAC 120

QY 121 AATATTCCG 130
DB 121 AATATTACAG 130

RESULT 4
AX047116
LOCUS AX047116 132 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 30 from Patent WO0068265.
ACCESSION AX047116
VERSION AX047116.1 GI:11876502
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1
AUTHORS Selsted,M.E., Tang,Y.Q., Yuan,J. and Ouellette,A.J.
TITLE Antimicrobial theta defensins and methods of using same
JOURNAL Patent: WO 0068265-A 30 16-NOV-2000;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA (US)
FEATURES Location/Qualifiers
source 1. .132
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic Construct"

ORIGIN

Query Match 82.7%; Score 109.2; DB 6; Length 132;
Best Local Similarity 90.0%; Pred. No. 2.7e-26;
Matches 117; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTGCCGACTTTTGGCGGAGAA 60
DB 1 GGACCTTGTCTACTAGTCCCTTACCGAGTAAGAAATGTACCAGACTTTTGGCGGAGAA 60

QY 61 GGCGAAAGTCTCAGTCCGCTCTCCGAACTCCACGAAACGCTGCTCTCTCAACGGTTGAC 120
DB 61 GGCGAAAGTCTCAGTCCGCTCTCCGAACTCCACGAAACGCTGCTCTCTCAACGGTTGAC 120

QY 121 AATATTCCG 130
DB 121 AATATTACAG 130

RESULT 5
AL844144/c
LOCUS AL844144 178965 bp DNA linear ROD 16-MAY-2003
DEFINITION Mouse DNA sequence from clone RP23-204D17 on chromosome 2, complete
sequence.
ACCESSION AL844144
VERSION AL844144.10 GI:30842726

KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Almeida, J.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT On May 17, 2003 this sequence version replaced gi:30140518. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submissions corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/Celegans/wormpep> RP23-204D17 is from the RPI-23 Mouse BAC Library constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6.
FEATURES
 source
 1..178965
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosomes="2"
 /clone="RP23-204D17"
 /clone_lib="RPI-23"
ORIGIN
 Query Match 24.7%; Score 32.6; DB 10; Length 178965;
 Best Local Similarity 58.9%; Pred. No. 9;
 Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
 QY 9 TCTACTAGTCCTTACCGAGTAGGAATGTCGGACATTTTCGGCGGAGGCGGAAG 68
 DB 117248 TGTACAAGTCCCTTAACAGAGGCTGCACACTGGACACGATGTGGAGGGAAGGTCAC 117189
 QY 69 TCTCAGTCGCTCTCCGAACTCCACGGGAACGTCGTG 103
 DB 117188 TCTCAGCCACTACGAAAGCTAGGCACTATCTG 117154
RESULT 6
AC107644/c
LOCUS AC107644
DEFINITION Mus musculus clone RP23-55K6, *** SEQUENCING IN PROGRESS ***, 14 ordered pieces.

ACCESSION AC107644
VERSION GI:30315786
KEYWORDS HTG; HTGS PHASE2; HTGS FULLTOP; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE 1 (bases 1 to 241189)
JOURNAL Mus musculus, clone RP23-55K6
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 241189)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckgalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J., Charaz, B., Choepel, J., Collangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lakocque, K., Lamazares, R., Landers, P., Lenoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
TITLE Direct Submission
JOURNAL Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 241189)
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On May 2, 2003 this sequence version replaced gi:22381169. All repeats were identified using RepeatMasker:
 Smit, A. F. A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information

 Center project name: L14267
 Center clone name: 55_K_6

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.
 * This sequence will be replaced.
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 2835: contig of 2835 bp in length
 * 2836: gap of 100 bp
 * 2537: contig of 22438 bp in length
 * 25374: gap of 100 bp
 * 25474: contig of 2675 bp in length
 * 28149: gap of 100 bp
 * 28249: contig of 16614 bp in length
 * 44863: gap of 100 bp
 * 44863: contig of 4997 bp in length
 * 49360: gap of 100 bp
 * 60823: contig of 10764 bp in length
 * 60824: gap of 100 bp
 * 60824: contig of 11628 bp in length
 * 72552: gap of 100 bp
 * 72552: contig of 13997 bp in length
 * 86648: gap of 100 bp
 * 86649: contig of 26261 bp in length
 * 86749: gap of 100 bp
 * 113010: contig of 3568 bp in length
 * 113110: gap of 100 bp
 * 116678: contig of 9597 bp in length
 * 126375: gap of 100 bp
 * 126375: contig of 40702 bp in length
 * 167177: gap of 100 bp
 * 167277: contig of 69977 bp in length
 * 237254: gap of 100 bp
 * 237354: contig of 3836 bp in length.

FEATURES

source

1..241189
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP23-55K6"
 /clone_lib="RPCI-23 Female Mouse BAC"

ORIGIN

Query Match 24.7%; Score 32.6; DB 2; Length 241189;
 Best Local Similarity 58.9%; Pred. No. 9.3; Indels 0; Gaps 0;
 Matches 56; Conservative 0; Mismatches 39;
 QY 9 TCTACTAGTCCCTTACCGATAGGAAATGTGCGGACATTTTTCGCGGAGAGGCGCAAG 68
 Db 27353 TGTACAGTCCCTTAACAGAGGCTGCACACTGACACAGATGTGGAGGAAAGGCTCAAC 27294
 QY 69 TCTCAGTGGCTTCCTCCGAACTCCACGAAACGTCGT 103
 Db 27293 TCTCAGCACTCAACGAAACGCTAGGCAAGTATCTG 27259

RESULT 7

AK129480/c 5339 bp mRNA linear ROD 21-NOV-2003
 LOCUS Mus musculus mRNA for mKIAA3015 protein.
 DEFINITION AK129480
 ACCESSION AK129480.1 GI:37360623
 VERSION FLI_CDNA.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE 1
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Koseki, H., Hiraoka, S.,
 Saga, Y., Nagase, T., Ohara, O. and Koga, H.

TITLE

Prediction of the coding sequences of mouse homologues of KIAA
 gene: III. the complete nucleotide sequences of 500 mouse
 KIAA-homologous cDNAs identified by screening of terminal sequences
 of cDNA clones randomly sampled from size-fractionated libraries
 DNA Res. 10 (4), 167-180 (2003)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..5339

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="mbg13024"

/tissue_type="brain"

/dev stage="adult"

/note="vector:modified pBC SK+"

1..5339

/gene="mKIAA3015"

<258..2834

/gene="mKIAA3015"

/note="CDS is predicted by in silico analysis. Start codon
 is not identified."

/codon_start=1

/evidence="not experimental"

/product="mKIAA3015 protein"

/protein_id="BAC98290.1"

/db_xref="GI:37360624"

/translation="PNPPEIAGAAGAGGAPVAERQLGSRVYRPANDKIEDINGCP
 KNRQMIENIDACLNLAAGINTOGLSABEIRGNLKAITGLFFSLRYKQOQOQQ
 QOPEKPLSSPLPPAGSOVAGAPSOQAGTPOHGLATPOAPCOLLPVSHOQGTQ
 VEQSSASSKDSQSKIIIRFLCQKISRLPGTARVAAGSEAKTEGSAARNRSS
 QSFNNYDKSPVTSPPPPPPSHREKEPLASSASHPGSENVPALSPSPVNCVS
 SSAPQPSMTSKPWSKSUSVKTATSMLSVKGAPGAPRPTPEAMKAPANNQKML
 EKLKFNKSGSAGSGASRDTSCELEILPSFEETELEATANEALSTVGASSP
 KIALKGIAQRTFSAITNKKSPKNEKEKEKEKEKEKEKEKEKEKEKEKEKEKE
 PDLKETKADLGSVAVTEPKSSKIASIPKGGKLNSTKKEATAPSHSGIPKPMKN
 VSAKSPAPIPPEGERSGKLSGLPPKQALDSRSHSSSSSLASSEKGPQGTSLN
 PSLSQTVSGVGTITGTSNTVSQVLPQOQYNHPNTATVAFPLYRSQTDPEGVNT
 AESSSAGVMEPSHYTKSGQPALEELTEDEPEARLRNTVKNIAOLRNLETMSLSGT
 QVHTLETTFDTNVTTEISGRSILSLTGRPTLSWRLGSSSPRLQAGDAPSMNGYP
 PRNARFISAEAGRYVYAPLRRLQASRSSICHDVSDKADDDVLDLGGISMDAPGY
 MSQGDVLSKNIIRSDDTISGTMTDGLGLYTRRLNRLPDGMVAVVRETQLRNTSLGLGDA
 DR"

ORIGIN

Query Match 23.0%; Score 30.4; DB 10; Length 5339;
 Best Local Similarity 55.8%; Pred. No. 35;
 Matches 58; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 QY 15 AGTCCCTTACCGATAGGAAATGTGCGGACATTTTTCGCGGAGAGGCGCAAGTCTCAG 74
 Db 3227 AGTTGAGTATAGAGAGGAGGAGTGGACAGTCTCTGGGCGCTGCAGAGAGAGTCCGAG 3168
 QY 75 TCGCTCTCCGAACCTCCACGGAACGCTGTCTCTCTCAAAACGGTTG 118
 Db 3167 TAACAGTCAAAACAGCAACCAACCGCTCAAGCCGCAACCCCTTG 3124

RESULT 8

AC120854

LOCUS

DEFINITION

Mus musculus clone RP23-329B6, WORKING DRAFT SEQUENCE, 12 unordered

207736 bp DNA linear HTG 09-MAR-2003


```

misc_feature 142259..207736
ORIGIN
Query Match 23.0%; Score 30.4; DB 2; Length 207736;
Best Local Similarity 55.8%; Pred. No. 52;
Matches 58; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 15 AGTCCTTACCGAGTAAGGAATGTGCGGAGCTTTTGGCGGAGAGGCGAAGTCTCAG 74
Dbb 36536 AGTTGAGTATGAGAGAGGAGTGTAGGCGAGGTCCTGCGGCGCTGCGAGAGAGTCTGAG 36595

QY 75 TCGCTTCGGAACCTCAGCAACGCTCTCTCTCTCAACGGTTG 118
Dbb 36596 TAAACGTCAAAACAGCAACCAACCGCTCAACGCCGAAACCTTG 36639

RESULT 9
AC123263 Rattus norvegicus clone CH230-296E17, *** SEQUENCING IN PROGRESS
LOCUS AC123263
DEFINITION Rattus norvegicus clone CH230-296E17, *** SEQUENCING IN PROGRESS
AC123263
VERSION HTG: HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 216422)
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gonzalez, E., Geet, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.I., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshew, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Nwaokelumen, O., Okwuonu, G., Olampunseagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.O.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
Sneed, A., Sodergrsten, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,

```

```

Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, S., Wlarczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G., and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 216422)
Worley, K.C.
Direct Submission
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 216422)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 14, 2002 this sequence version replaced gi:21671637.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTUG
Center clone name: CH230-296E17
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 186479 bases at least Q40
Consensus quality: 187629 bases at least Q30
Consensus quality: 188191 bases at least Q20
Estimated insert size: 199916; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 213497: contig of 213497 bp in length
* 213498 213597: gap of unknown length
* 213598 214695: contig of 1098 bp in length
* 214696 214795: gap of unknown length
* 214796 216422: contig of 1627 bp in length.
* Location/Qualifiers
* 1. 216422
* /organism="Rattus norvegicus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-296E17"
* 206750..207923
* /note="wgs_contig"
misc_feature

```


QY 83 CGAAGTCCAGGAACGCTGCTCCTCAAC 113
 |||||
 Db 76234 AGATTACAGAACTTGAGATTTTAAAC 76264
 |||||

RESULT 12
 BX248984 191884 bp DNA linear ROD 17-JUN-2003
 LOCUS Mouse DNA sequence from clone RP23-278P12 on chromosome 1, complete
 DEFINITION sequence.
 ACCESSION BX248984
 VERSION EX248984.11 GI:31873538
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 JOURNAL 1 (bases 1 to 191884)
 DIRECT SUBMISSION
 COMMENT humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Jun 17, 2003 this sequence version replaced gi:30350055.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to confirm this sequence. Sequence data
 from the whole genome shotgun alone has only been used where it has
 a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information
 on the WORMPEP database can be found at
<http://www.sanger.ac.uk/Projects/C.elegans/wormpep> RP23-278P12 is
 from the RCI-23 Mouse BAC Library
 constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACE3.6
 Location/Qualifiers
 1. 191884
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="1"
 /clone="RP23-278P12"
 /clone_lib="RP23-278P12"

FEATURES
 source

ORIGIN
 Query Match 22.9%; Score 30.2; DB 10; Length 191884;
 Best Local Similarity 58.2%; Pred. No. 60;
 Matches 53; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 11 TACTAGTCCCTACGAGTAAGGAATTTAGCCGACTTTGGCGAGAGCGGAAGTC 70
 |||||
 |||||

Db 41372 TAGTTATCTCTCAGTGAATTAAGGAATTTATACCAAGATTTTAGCAGCATCATGCAGAAACAC 41431
 QY 71 TCAGTCGCTCTCCGAAGTCCACGGAACGTC 101
 |||||
 Db 41432 TTAGTATCTCTATCTACCAAGTCTTTAAATTC 41462
 |||||

RESULT 13
 AC129542 194807 bp DNA linear ROD 26-SEP-2003
 LOCUS Mus musculus chromosome 1, clone RP23-249L15, complete sequence.
 DEFINITION Mus musculus chromosome 1, clone RP23-249L15, complete sequence.
 ACCESSION AC129542
 VERSION AC129542.9 GI:34610344
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 JOURNAL 1 (bases 1 to 194807)
 DIRECT SUBMISSION
 COMMENT Mus musculus chromosome 1, clone RP23-249L15
 Unpublished
 2 (bases 1 to 194807)
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
 Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
 Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Minova, T., Mienga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
 Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 194807)
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,
 Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
 Ferrira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrim, J., Meneus, L., Minova, T., Mienga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (24-JUL-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 194807)
 Birren, B., Nusbaum, C. and Lander, E.
 Direct Submission
 Submitted (13-SEP-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

5 (bases 1 to 194807)
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
 Boguslavskiy,L., Boukhgalter,B., Camarata,B., Chang,J., Choepel,Y.,
 Collamore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fato,S.,
 Ferreira,P., Flitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kanat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
 Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., MacLean,C.,
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
 Meldrum,J., Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
 Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
 Vassiliiev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (26-SEP-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 13, 2003 this sequence version replaced gi:32328999.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L26921
 Center clone name: 249_L15
 ----- Location/Qualifiers
 1..194807
 organism="Mus musculus"
 mol_type="genomic DNA"
 db_xref="taxon:10090"
 chromosome="1"
 map="1"
 clone="RP23-249L15"
 clone_lib="RPCI-23 Female Mouse BAC"
 note="clone boundary"
 misc_feature
 1..6
 note="PCR product sequence only"
 repeat_region
 complement(68..283)
 /rpt_family="L1_MM"
 unsure
 111..115
 note="pCR product sequence only"
 repeat_region
 364..600
 /rpt_family="Lx2"
 repeat_region
 601..995
 /rpt_family="MTA"
 repeat_region
 996..1631
 /rpt_family="Lx2"
 repeat_region
 1860..1952
 /rpt_family="(TA)n"
 repeat_region
 complement(2224..2904)
 /rpt_family="Lx"
 repeat_region
 3271..3325
 /rpt_family="AT_rich"
 repeat_region
 4558..4584
 /rpt_family="AT_rich"
 repeat_region
 5413..5465
 /rpt_family="(TTAA)n"
 repeat_region
 complement(5491..5788)
 /rpt_family="Lx5"
 repeat_region
 6490..6679

Query Match 22.9%; Score 30.2; DB 10; Length 194807;
 Best Local Similarity 58.2%; Pred. No. 60;
 Matches 53; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

| | | | | |
|----|--------|--|-------|----|
| QY | 11 | TACTAGTCCCTTACCGAGTAGGAATACTGCCGACTTTTCGGCAGAACGCAAGGC | AAGTC | 70 |
| Db | 124982 | TAGTTATTGTCTACTGTATAGGANTTATACCAGATTTTAGCAGCATCATGCGAACAAC | | |
| QY | 71 | TCAGTCGCTCTCCGAATCCACGGAACGTC | 101 | |

| TITLE | Direct | Submission |
|-------|--------|---|
| D5 | 135862 | TTAGAACAGTGGCTGTCTTTGGCCCTGAAGCAGGGCAGGCTTAGCTCGCCCTCACACCT |

RESULT 15

AC099298

LOCUS

DEFINITION

AC099298 239140 bp DNA linear HTG 13-MAY-2003
Rattus norvegicus clone CH230-40116, WORKING DRAFT SEQUENCE, 5
unordered pieces.

ACCESSION

AC099298

VERSION

HTG; HTG_PHASE1; HTG_DRAFT; HTG_FULLTOP.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

Rattus norvegicus

ORGANISM

Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;

Rattus

1 (bases 1 to 239140)

REFERENCE

AUTHORS

Muzny, D., Maric, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Evans, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flegge, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gueorge, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Hulek, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpathy, S., Kelly, S., Khan, Z., King, D., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Loreschweta, L., Loulseg, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Maindarte, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Napua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, F.Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Riley, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Speed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Streimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 239140)

REFERENCE

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:25085778.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GLEN

Center clone name: CH230-40116

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 225133 bases at least Q40

Consensus quality: 227827 bases at least Q30

Consensus quality: 230020 bases at least Q20

Estimated insert size: 239353; sum-of-contigs estimation

Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 5 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

*-----

* 1 64307: contig of 64307 bp in length

* 64308 64407: gap of unknown length

* 64408 223476: contig of 159069 bp in length

* 223477 223576: gap of unknown length

* 223577 227284: contig of 3708 bp in length

* 227285 227384: gap of unknown length

* 227385 231763: contig of 4379 bp in length

* 231764 231863: gap of unknown length

* 231864 239140: contig of 7277 bp in length.

*----- Location/Qualifiers

1. 239140

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-40116"

1. 1306

/note="wgs contig"

3687. 6304

/note="wgs contig"

6355. 8173

/note="wgs contig"

223577. 225471

/note="wgs contig"

ORIGIN

Query Match

Best Local Similarity

Matches

54; Conservative

0; Mismatches

40; Indels

0; Gaps

22.7%; Score 30; DB 2; Length 239140;

57.4%; Pred. No. 72;

0; Mismatches

40; Indels

0; Gaps

0; Mismatches

0; Indels

0; Gaps

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

2 (bases 1 to 239140)

REFERENCE

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

REFERENCE

QY 29 TAAGGAAATGTCCGACCTTTTGGCGCGAGAAAGTCTCAGTCGCTCTCCGAACT 88
Db 45940 TTAGAACAGGTGGCTGTCTTTGCCCTGAAGCAGGCGACGGTCTAGCTCGCCCTCACACCT 45999
QY 89 CCACGGAAACGTCTGCTCTCAACGGTTGACAA 122
Db 46000 CCATCTCATTTGACCGCTCCTTGCCCTGGCTGGCAA 46033

Search completed: October 11, 2004, 10:08:24
Job time : 1129.5 secs

XX SQ Sequence 2817 BP; 895 A; 722 C; 746 G; 464 T; 0 U; 0 Other;
Query Match 22.4%; Score 29.6; DB 4; Length 2817;
Best Local Similarity 61.8%; Pred. No. 1.8;
Matches 47; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 1 GCACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTCGGGACTTTTGGCGGAGAA 60
DB 1524 GCGCCGCGAAAAGTTGGCGCTGTACGTGTACGAATATCTGTGCACGTGGCGCCAGAA 1583
QY 61 GCGCAAGTCTCAGTC 76
DB 1584 GCGGCACAGACATTC 1599
RESULT 4
ID ABL12502/c
XX ABL12502 standard; cDNA; 6398 BP.
AC ABL12502;
XX
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 31988.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEXE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX P-PSDB; ABB68399.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX Claim 1; SEQ ID NO 31988; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 6398 BP; 1456 A; 1470 C; 1528 G; 1944 T; 0 U; 0 Other;
Query Match 22.4%; Score 29.6; DB 4; Length 6398;
Best Local Similarity 61.8%; Pred. No. 2.5;
Matches 47; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 1 GCACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTCGGGACTTTTGGCGGAGAA 60
DB 2294 GCGCCGCGAAAAGTTGGCGCTGTACGTGTACGAATATCTGTGCACGTGGCGCCAGAA 2235

QY 61 GCGCAAGTCTCAGTC 76
DB 2234 GCGGCACAGACATTC 2219
RESULT 5
ABX09140/c
ID ABX09140 standard; DNA; 82993 BP.
XX
AC ABX09140;
XX
DT 08-APR-2003 (first entry)
DE Mycobacterium tuberculosis H37Rv BAC clone BAC-Rv221.
XX
XX Mycobacteriosis; survival; virulence; protective antigen; vaccine;
KW mycobacterial disease; tuberculosis; leprosy; ds; cosmid.
XX
XX Mycobacterium tuberculosis.
XX
XX WO200274903-A2.
XX
XX 26-SEP-2002.
XX
XX 22-FEB-2002; 2002WO-IB001-973.
XX
XX 22-FEB-2001; 2001US-0270123P.
XX
XX (INSP) INST PASTEUR.
XX
XX Cole S;
XX
XX WPI; 2002-759885/82.
XX
XX Identifying and selecting genes for survival or virulence of mycobacteria
XX by a comparative genomic analysis of the sequences of Mycobacterium
XX tuberculosis and M. leprae.
XX
XX Disclosure; Fig 4; 874pp; English.
XX
XX This invention relates to a novel method for identifying essential genes
XX for survival or virulence of mycobacteria species. The method comprises
XX aligning the genomic sequence of a first mycobacterium species on a
XX genomic sequence of a second mycobacterium species and selecting a
XX polynucleotide sequence that is highly conserved in both genomes with no
XX counterparts in other bacterial genomic sequences and that corresponds to
XX an essential gene for the survival or virulence of mycobacterium species.
XX The method of the invention is useful for detecting M. tuberculosis or M.
XX leprae infection. The method reduces the number of potential new targets
XX and protective antigens for new drugs and vaccine compositions to treat
XX and prevent mycobacterial diseases, particularly tuberculosis and
XX leprosy. The present sequence represents a Mycobacterial cosmid DNA
XX sequence used in the method of the invention
XX
SQ Sequence 82993 BP; 14981 A; 26628 C; 26719 G; 14765 T; 0 U; 0 Other;
Query Match 22.1%; Score 29.2; DB 6; Length 82993;
Best Local Similarity 57.8%; Pred. No. 10;
Matches 52; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 5 CTGCTCTACTAGTCCCTTACCGAGTAAGAAATGTCGGGACTTTTGGCGGAGAGGCG 64
DB 79038 CTCGATGACAGTCAGCAAACTACCGAGTAATCCGAGCCGAGTCTTGACCAACGATTCG 78979
QY 65 AAAGTCTCAGTCGCTCTCCGAATCCACGCG 94
DB 78978 AACGGCCCGTGGTCAACGAACAAGACGG 78949
RESULT 6
AAI99682 21/c
Continuation (22 of 45) of AAI99682 from base 2100001 (Mycobacterium tuberculosis strain

PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
DR P-PSDB; AAG92650.
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX Claim 8; SEQ ID NO 2904; 246pp + Sequence Listing; English.
XX The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Corynebacterium, and identifying a homologue of a gene derived from
CC Corynebacterium. Corynebacterium bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX Sequence 1983 BP; 410 A; 508 C; 558 G; 507 T; 0 U; 0 Other;
SQ Query Match 21.8%; Score 28.8; DB 5; Length 1983;
Best Local Similarity 58.0%; Pred. No. 3.2;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 25 CGAGTAAGAAATGTCCGACTTTTGGCGGAGGAGGAAAGTCTCAGTGGCTCCG 84
DB 18 CCGCAACACATCTCGCGACATTTGGCGGAGGAGGAAACATTTGCGCGGCACTG 77
QY 85 AACTCCACGAAACGTCTCTCTCTCAAA 112
DB 78 TGCACGCGTTTACGCTCTCTCTCAAA 105
RESULT 9
AAS81289
ID AAS81289 standard; cDNA; 2012 BP.
XX AAS81289;
AC AAS81289;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #17093.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US0008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
XX 11-JAN-2001.

DR WPI; 2001-639362/73.
XX P-PSDB; ABG17102.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 17093; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 2012 BP; 406 A; 538 C; 658 G; 409 T; 0 U; 1 Other;
SQ Query Match 21.8%; Score 28.8; DB 5; Length 2012;
Best Local Similarity 62.5%; Pred. No. 3.2;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 38 GTGCGGACATTTTGGCGGAGGAGGAAAGTCTCAGTCTCCGAACTCCACGAAA 97
DB 1699 GTGCGGCTTTTCTCTTGGAGGCCAAGGCTGGCTTGGACTTCGAGCTGACTCACT 1758
QY 98 CGTCTGCTCTC 109
DB 1759 CATCTCCCTTC 1770
RESULT 10
AAF32543
ID AAF32543 standard; DNA; 5969 BP.
XX AAF32543;
AC AAF32543;
XX 11-SEP-2003 (revised)
DT 20-APR-2001 (first entry)
XX Brevibacterium lactofermentum sucrose PTS enzyme II DNA SEQ ID NO:1.
DE Brevibacterium lactofermentum; sucrose PTS enzyme II; sucrose;
KW phosphoenolpyruvate:carbohydrate phosphotransferase system; glucose;
KW coryneform bacterium; phosphoenolpyruvate-sugar transport system; ds.
XX Corynebacterium glutamicum.
OS Key Location/Qualifiers
FH CDS 3779..5764
FT /*tag= a
FT /product= "sucrose PTS enzyme II"
XX WO200102584-A1.
XX 11-JAN-2001.

PF 30-JUN-2000; 2000WO-JF004348.
XX
PR 02-JUL-1999; 99JP-00189512.
XX
PA (AJIN) AJINOMOTO CO INC.
XX
PI Izui M, Sugimoto M, Nakamatsu T, Kurahashi O;
XX
DR WPI; 2001-138150/14.
DR P-PSDB; AAB69080.
XX
XX Gene encoding sucrose phosphoenolpyruvate-sugar transport system enzyme
PT II obtained by cassette ligation-mediated amplification of downstream
PT domain of corynebacterium bacterium sucrose gene, with sucrose-binding
PT activity.
XX
PS Claim 3; Page 22-29; 45pp; Japanese.
XX
CC The present sequence encodes the Brevibacterium lactofermentum sucrose
CC PTS (phosphoenolpyruvate:carbohydrate phosphotransferase system or
CC phosphoenolpyruvate-sugar transport system) enzyme II, which has sucrose-
CC binding activity. A corynebacterium bacterium produced with the sucrose PTS
CC enzyme II gene can have more efficient sugar uptake, and improved amino-
CC acid and nucleic acid productivity. The sucrose PTS gene and it's
CC disrupted gene, such as one without the sucrose PTS function, can be used
CC to produce new breeds of corynebacterium bacterial strains to uptake sugar
CC more efficiently e.g. glucose only or and sucrose, and can have improved
CC amino-acid and nucleic acid productivity. (Updated on 11-SEP-2003 to
CC standardise OS field)
XX
SQ Sequence 5969 BP; 1388 A; 1575 C; 1578 G; 1428 T; 0 U; 0 Other;
Query Match 21.8%; Score 28.8; DB 4; Length 5969;
Best Local Similarity 58.0%; Pred. No. 4.9;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 25 CGAGTAAGGAATGTGCGGACATTTTCGCGGAGAGCGAAGTCTCAGTCTCTCCG 84
Db 3796 CGCGCAACGCATCTCGGACATTTTCGCGGAGAGCAACATTTGCGCGCGCACACTG 3855
QY 85 AACTCCACGAAACGTCTGCTCTCAA 112
Db 3856 TGCAACGGTTTACGCTCTGCTCAA 3883
RESULT 11
AAH68533/c
ID AAH68533 standard; DNA; 349980 BP.
XX
AC AAH68533;
XX
XX 26-SEP-2001 (first entry)
DT
XX C glutamicum coding sequence fragment SEQ ID NO: 7068.
DE
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; ds.
XX
XX Corynebacterium glutamicum.
OS
XX BP1108790-A2.
XX
XX 20-JUN-2001.
PD
XX 18-DEC-2000; 2000BP-00127688.
XX
XX 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
XX
PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
PS Disclosure; SEQ ID NO 7068; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium bacterium, measuring expression amount and analyzing
CC the expression profile or expression pattern of a gene derived from
CC corynebacterium bacterium, and identifying a homologue of a gene derived from
CC corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 349980 BP; 81250 A; 97718 C; 90621 G; 80391 T; 0 U; 0 Other;
Query Match 21.8%; Score 28.8; DB 5; Length 349980;
Best Local Similarity 58.0%; Pred. No. 26;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 25 CGAGTAAGGAATGTGCGGACATTTTCGCGGAGAGCGAAGTCTCAGTCTCTCCG 84
Db 111789 CGCGCAACGCATCTCGGACATTTTCGCGGAGAGCAACATTTGCGCGCGCACACTG 111730
QY 85 AACTCCACGAAACGTCTGCTCTCAA 112
Db 111729 TGCAACGGTTTACGCTCTGCTCAA 111702
RESULT 12
AAI99682.12
Continuation (13 of 45) of AAI99682 from base 1200001 (Mycobacterium tuberculosis strain
WP Sequence split into 45 fragments LOCUS AAI99682 Accession AAI99682
WP Fragment Name Begin End
WP AAI99682_00 1 110000
WP AAI99682_01 100001 210000
WP AAI99682_02 200001 310000
WP AAI99682_03 300001 410000
WP AAI99682_04 400001 510000
WP AAI99682_05 500001 610000
WP AAI99682_06 600001 710000
WP AAI99682_07 700001 810000
WP AAI99682_08 800001 910000
WP AAI99682_09 900001 1010000
WP AAI99682_10 1000001 1110000
WP AAI99682_11 1100001 1210000
WP AAI99682_12 1200001 1310000
WP AAI99682_13 1300001 1410000
WP AAI99682_14 1400001 1510000
WP AAI99682_15 1500001 1610000
WP AAI99682_16 1600001 1710000
WP AAI99682_17 1700001 1810000
WP AAI99682_18 1800001 1910000
WP AAI99682_19 1900001 2010000
WP AAI99682_20 2000001 2110000
WP AAI99682_21 2100001 2210000
WP AAI99682_22 2200001 2310000
WP AAI99682_23 2300001 2410000
WP AAI99682_24 2400001 2510000
WP AAI99682_25 2500001 2610000
WP AAI99682_26 2600001 2710000
WP AAI99682_27 2700001 2810000
WP AAI99682_28 2800001 2910000
WP AAI99682_29 2900001 3010000

WP AAI99682_30 3000001 3110000
WP AAI99682_31 3100001 3210000
WP AAI99682_32 3200001 3310000
WP AAI99682_33 3300001 3410000
WP AAI99682_34 3400001 3510000
WP AAI99682_35 3500001 3610000
WP AAI99682_36 3600001 3710000
WP AAI99682_37 3700001 3810000
WP AAI99682_38 3800001 3910000
WP AAI99682_39 3900001 4010000
WP AAI99682_40 4000001 4110000
WP AAI99682_41 4100001 4210000
WP AAI99682_42 4200001 4310000
WP AAI99682_43 4300001 4410000
WP AAI99682_44 4400001 4411529

Query Match 21.5%; Score 28.4; DB 4; Length 110000;
Best Local Similarity 54.9%; Pred. No. 23;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 2 GACCTGCTCTACTAGTCCCTTACCGAGTAAGAAATGTCCGGACTTTTGGCGGAGAAAG 61
|||||
Db 46078 GACTTCGCCACAAAGCGGTGTCGGCATGCGCAAGCAGTTTCGGCGGACACGCGGAGAAA 46137
|||||

QY 62 GCGAAGTCTCAGTCCCTCCGACTCCACGGAACGCTCG 103
|||||

Db 46138 CCGGTAAGTCTAGTCCCTGACGAAGTCCACACGACGTCGG 46179
|||||

RESULT 13
AAI99683_12
Continuation (13 of 44) of AAI99683 from base 1200001 (Mycobacterium tuberculosis strain
WP Sequence split into 44 fragments LOCUS AAI99683 Accession AAI99683
Fragment Name Begin End
WP AAI99683_00 1 110000
WP AAI99683_01 100001 210000
WP AAI99683_02 200001 310000
WP AAI99683_03 300001 410000
WP AAI99683_04 400001 510000
WP AAI99683_05 500001 610000
WP AAI99683_06 600001 710000
WP AAI99683_07 700001 810000
WP AAI99683_08 800001 910000
WP AAI99683_09 900001 1010000
WP AAI99683_10 1000001 1110000
WP AAI99683_11 1100001 1210000
WP AAI99683_12 1200001 1310000
WP AAI99683_13 1300001 1410000
WP AAI99683_14 1400001 1510000
WP AAI99683_15 1500001 1610000
WP AAI99683_16 1600001 1710000
WP AAI99683_17 1700001 1810000
WP AAI99683_18 1800001 1910000
WP AAI99683_19 1900001 2010000
WP AAI99683_20 2000001 2110000
WP AAI99683_21 2100001 2210000
WP AAI99683_22 2200001 2310000
WP AAI99683_23 2300001 2410000
WP AAI99683_24 2400001 2510000
WP AAI99683_25 2500001 2610000
WP AAI99683_26 2600001 2710000
WP AAI99683_27 2700001 2810000
WP AAI99683_28 2800001 2910000
WP AAI99683_29 2900001 3010000
WP AAI99683_30 3000001 3110000
WP AAI99683_31 3100001 3210000
WP AAI99683_32 3200001 3310000
WP AAI99683_33 3300001 3410000
WP AAI99683_34 3400001 3510000
WP AAI99683_35 3500001 3610000
WP AAI99683_36 3600001 3710000
WP AAI99683_37 3700001 3810000
WP AAI99683_38 3800001 3910000

WP AAI99683_39 3900001 4010000
WP AAI99683_40 4000001 4110000
WP AAI99683_41 4100001 4210000
WP AAI99683_42 4200001 4310000
WP AAI99683_43 4300001 4403765

Query Match 21.5%; Score 28.4; DB 4; Length 110000;
Best Local Similarity 54.9%; Pred. No. 23;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 2 GACCTGCTCTACTAGTCCCTTACCGAGTAAGAAATGTCCGGACTTTTGGCGGAGAAAG 61
|||||
Db 45607 GACTTCGCCACAAAGCGGTGTCGGCATGCGCAAGCAGTTTCGGCGGACACGCGGAGAAA 45666
|||||

QY 62 GCGAAGTCTCAGTCCCTCCGAACTCCACGGAACGCTCG 103
|||||

Db 45667 CCGGCTAACTAAGTCGCTGACGAAGTCCACACGACGTCGG 45708
|||||

RESULT 14
AAAS7511
ID AAAS7511 standard; DNA; 5271 BP.
XX
AC AAAS7511;
XX
DT 20-OCT-2000 (first entry)
XX
DE A TIGR (trabecular meshwork inducible glucocorticoid receptor) promoter.
KW TIGR; trabecular meshwork inducible glucocorticoid receptor; promoter;
KW Glaucoma; steroid sensitivity; progressive ocular hypertension;
KW vision loss; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT mutation replace(4256, G)
FT /*tag= d
FT /note= "TIGRmt4 mutant"
FT mutation replace(4337, G)
FT /*tag= a
FT /note= "TIGRmt1 mutant"
FT mutation replace(4950, T)
FT /*tag= b
FT /note= "TIGRmt2 mutant"
FT mutation 4998
FT /*tag= C
FT /note= "GNGT added to produce TIGRmt3 mutant"
FT mutation replace(5113, C)
FT /*tag= e
FT /note= "TIGRmt11 mutant"
XX
DN WO200042220-A1.
XX
PD 20-JUL-2000.
XX
PP 11-JAN-2000; 2000WO-US000559.
XX
PP 11-JAN-1999; 99US-00227881.
PP 07-MAY-1999; 99US-00306828.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Nguyen TD, Polansky JR, Chen F, Chen H;
XX WPI; 2000-491060/43.
XX
XX
XX
PT Diagnosis, prognosis and treatment of glaucoma, based on detecting
PT specific polymorphisms in the promoter of the trabecular meshwork
PT inducible glucocorticoid receptor gene.
PS Claim 79; Page 117-119; 122pp; English.
XX

CC The present sequence represents a TIGR (trabecular meshwork inducible
 CC glucocorticoid receptor) promoter, isolated from an individual without
 CC glaucoma. The specification describes a method for the diagnosis,
 CC prognosis and treatment of glaucoma, based on detecting specific
 CC polymorphisms in the promoter of the TIGR gene. The method is used for
 CC diagnosis and prognosis of glaucoma (of all types), steroid sensitivity
 CC and progressive ocular hypertension that leads to loss of vision.
 CC Glaucoma can be treated by administering an agent that binds to cis-
 CC acting elements within the TIGR promoter. The TIGR promoter (or other
 CC regulatory regions) can be used to express homologous or heterologous
 CC genes, particularly for tissue-specific expression of therapeutic
 CC transgenes for treating glaucoma, also to generate transgenic animals and
 CC in screening for compounds (specific modulators) with diagnostic or
 CC therapeutic potential. Fragments of the TIGR sequence can be used as
 CC amplification primers or probes, e.g. for isolating related sequences in
 CC non-human animals
 XX
 SQ Sequence 5271 BP; 1476 A; 1138 C; 1231 G; 1426 T; 0 U; 0 Other;

Query Match 20.9%; Score 27.6; DB 3; Length 5271;
 Best Local Similarity 56.7%; Pred. No. 13;
 Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
 QY 21 TTACCGAGTAAGGAAATGTCGGGACTTTTGGCGGAGAGGCGGAAAGTCTCAGTCGGTC 80
 DB 2040 TTATTGAGTACTTATATCTGCCAGACACGAGACAAATGTTGACCAAGCGTCACTG 2099
 QY 81 TCCGAACCTCCACGGAACGTCCTCTCTCA 110
 DB 2100 CCTACCTTCGTGGAGGTGACAGTTTCTCA 2129

RESULT 15
 AAV51361
 ID AAV51361 standard; DNA; 5299 BP.
 XX
 AC AAV51361;
 XX
 DT 27-OCT-1998 (first entry)
 XX
 DE Human TIGR promoter region DNA.
 XX
 KW TIGR; trabecular meshwork induced glucocorticoid response protein; human;
 KW diagnosis; glaucoma; polymorphism; steroid sensitivity; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9832850-A1.
 XX
 PD 30-JUL-1998.
 XX
 PF 09-JAN-1998; 98WO-US000468.
 XX
 PR 28-JAN-1997; 97US-00791154.
 XX
 PR 26-SEP-1997; 97US-00938669.
 XX
 PR (REGC) UNIV CALIFORNIA.
 XX
 PI Nguyen TD, Polansky JR, Chen P, Chen H;
 XX
 DR WPI; 1998-427946/36.
 XX
 PT Use of TIGR nucleic acid sequences - used for, e.g. developing products
 PT for diagnosis, prognosis and treatment of glaucoma.
 XX
 PS Claim 34; Fig 1; 105pp; English.
 XX

CC This sequence is a trabecular meshwork induced glucocorticoid response
 CC protein (TIGR) promoter region which is used in a method for diagnosing
 CC glaucoma in a patient. The method involves the detection of polymorphisms
 CC whose presence is predictive of a mutation affecting TIGR response in the
 CC patient and can be diagnostic of glaucoma or steroid sensitivity. Base
 CC substitutions and base additions upstream of and within TIGR exons can

CC also be used to diagnose glaucoma
 XX
 SQ Sequence 5299 BP; 1482 A; 1151 C; 1235 G; 1431 T; 0 U; 0 Other;
 Query Match 20.9%; Score 27.6; DB 2; Length 5299;
 Best Local Similarity 56.7%; Pred. No. 13;
 Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
 QY 21 TTACCGAGTAAGGAAATGTCGGGACTTTTGGCGGAGAGGCGGAAAGTCTCAGTCGGTC 80
 DB 2040 TTATTGAGTACTTATATCTGCCAGACACGAGACAAATGTTGACCAAGCGTCACTG 2099
 QY 81 TCCGAACCTCCACGGAACGTCCTCTCTCA 110
 DB 2100 CCTACCTTCGTGGAGGTGACAGTTTCTCA 2129
 Search completed: October 11, 2004, 09:30:40
 Job time : 191 secs

| | | | | | | | |
|---|----|------|------|---------|---|---------------------|-------------------|
| C | 28 | 26 | 19.7 | 2936 | 3 | US-08-943-731-64 | Sequence 64, Appl |
| C | 29 | 26 | 19.7 | 38682 | 3 | US-08-943-731-2 | Sequence 2, Appl |
| C | 30 | 25.2 | 19.1 | 1288 | 2 | US-09-172-977-2 | Sequence 2, Appl |
| C | 31 | 25.2 | 19.1 | 1288 | 4 | US-09-404-108-2 | Sequence 2, Appl |
| C | 32 | 25 | 18.9 | 1220 | 4 | US-09-721-870-9 | Sequence 9, Appl |
| C | 33 | 25 | 18.9 | 1985 | 4 | US-09-205-258-97 | Sequence 97, Appl |
| C | 34 | 24.8 | 18.6 | 546 | 4 | US-09-134-000C-2087 | Sequence 2087, Ap |
| C | 35 | 24.6 | 18.6 | 505 | 4 | US-09-621-976-15639 | Sequence 15639, A |
| C | 36 | 24.4 | 18.5 | 2148 | 4 | US-09-489-039A-3539 | Sequence 3539, Ap |
| C | 37 | 24.2 | 18.3 | 1587 | 4 | US-09-107-532A-984 | Sequence 984, App |
| C | 38 | 24.2 | 18.3 | 3149 | 4 | US-09-392-714-3 | Sequence 3, Appl |
| C | 39 | 24.2 | 18.3 | 3973 | 2 | US-08-602-093-6 | Sequence 6, Appl |
| C | 40 | 24.2 | 18.3 | 26684 | 4 | US-09-564-805-28 | Sequence 28, Appl |
| C | 41 | 24.2 | 18.3 | 41171 | 4 | US-08-311-731A-122 | Sequence 122, App |
| C | 42 | 24.2 | 18.3 | 1830121 | 4 | US-09-557-884-1 | Sequence 1, Appl |
| C | 43 | 24.2 | 18.3 | 1830121 | 4 | US-09-643-990A-1 | Sequence 1, Appl |
| C | 44 | 24 | 18.2 | 411 | 4 | US-09-252-991A-8838 | Sequence 8838, Ap |
| C | 45 | 24 | 18.2 | 732 | 4 | US-09-107-532A-456 | Sequence 456, App |

ALIGNMENTS

RESULT 1
US-09-309-487-31
; Sequence 31, Application US/09309487
; Patent No. 635318
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/309,487
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-09-309-487-31

| | | | | | |
|-----------------------|-----|--|--------------------|-----------|-------------|
| Query Match | | 100.0%; | Score 132; | DB 4; | Length 132; |
| Best Local Similarity | | 100.0%; | Pred. No. 2.1e-39; | | |
| Matches 132; | | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | GGACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTCCCGACTTTTGGCGCGAGAA | 60 | | |
| | 1 | GGACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTCCCGACTTTTGGCGCGAGAA | 60 | | |
| Db | 61 | GGCGAAAGTCTCAGTCCGCTCTCCGAACTCCACGAAACGTTCTCTCTCAACCGTTGAC | 120 | | |
| | 61 | GGCGAAAGTCTCAGTCCGCTCTCCGAACTCCACGAAACGTTCTCTCTCAACCGTTGAC | 120 | | |
| Qy | 121 | AATATTTCCGCA | 132 | | |
| | 121 | AATATTTCCGCA | 132 | | |
| Db | 121 | AATATTTCCGCA | 132 | | |
| | 121 | AATATTTCCGCA | 132 | | |

RESULT 2
US-09-967-808-31
; Sequence 31, Application US/09967808
; Patent No. 6514727
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 11, 2004, 09:19:46 ; Search time 39.5 Seconds
(without alignments)
1854.519 Million cell updates/sec

Title: US-10-009-317A-31
Perfect score: 132
Sequence: 1 ggacctctctactagtcc.....cgggtgacaatttccgca 132

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/FACTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|---------|-------|--------------------|
| 1 | 132 | 100.0 | 132 | 4 | US-09-309-487-31 |
| 2 | 132 | 100.0 | 132 | 4 | US-09-967-808-31 |
| 3 | 109.2 | 82.7 | 132 | 4 | US-09-309-487-30 |
| 4 | 109.2 | 82.7 | 132 | 4 | US-09-967-808-30 |
| C 5 | 29.2 | 22.1 | 4403765 | 3 | US-09-103-840A-2 |
| C 6 | 29.2 | 22.1 | 4411529 | 3 | US-09-103-840A-1 |
| 7 | 28.4 | 21.5 | 4403765 | 3 | US-09-103-840A-2 |
| 8 | 28.4 | 21.5 | 4411529 | 3 | US-09-103-840A-1 |
| 9 | 27.6 | 20.9 | 5271 | 4 | US-09-306-828-34 |
| 10 | 27.6 | 20.9 | 5300 | 3 | US-08-938-669A-1 |
| 11 | 27.6 | 20.9 | 5300 | 4 | US-09-306-828-1 |
| 12 | 27.6 | 20.9 | 5304 | 3 | US-08-938-669A-2 |
| 13 | 27.6 | 20.9 | 5304 | 4 | US-09-306-828-2 |
| 14 | 27.6 | 20.9 | 6169 | 3 | US-08-938-669A-3 |
| 15 | 27.6 | 20.9 | 6169 | 4 | US-09-306-828-3 |
| 16 | 27 | 20.5 | 612 | 4 | US-09-620-312D-529 |
| 17 | 26.2 | 19.8 | 1999 | 2 | US-08-645-900A-2 |
| 18 | 26.2 | 19.8 | 1999 | 2 | US-08-882-238A-2 |
| 19 | 26.2 | 19.8 | 1999 | 2 | US-08-667-790A-2 |
| 20 | 26.2 | 19.8 | 1999 | 2 | US-09-220-459-2 |
| 21 | 26.2 | 19.8 | 1999 | 3 | US-08-546-568B-2 |
| 22 | 26.2 | 19.8 | 2000 | 4 | US-09-056-285A-3 |
| 23 | 26.2 | 19.8 | 2099 | 3 | US-08-938-669A-5 |
| 24 | 26.2 | 19.8 | 2099 | 4 | US-09-306-828-5 |
| 25 | 26.2 | 19.8 | 2166 | 3 | US-08-822-999-1 |
| 26 | 26.2 | 19.8 | 43804 | 4 | US-09-171-461-1 |
| 27 | 26.2 | 19.8 | 1230025 | 4 | US-09-198-452A-1 |

```
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
; FILE REFERENCE: Same
; CURRENT APPLICATION NUMBER: US/09/967,808
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-09-967-808-31

Query Match      100.0%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 2.1e-39;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGGAAATGTCCGGACTTTTGGCGGAGAA 60
   |||||
DB 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGGAAATGTCCGGACTTTTGGCGGAGAA 60
   |||||

QY 61 GCGAAGAGTCTAGTCTCGTCTCCGAACTCCAGGAAACGTCTGCTCCTCAAAACGGTTGAC 120
   |||||
DB 61 GCGAAGAGTCTAGTCTCGTCTCCGAACTCCAGGAAACGTCTGCTCCTCAAAACGGTTGAC 120
   |||||

QY 121 AATATTCCGCA 132
   |||||
DB 121 AATATTCCGCA 132
   |||||

RESULT 3
US-09-309-487-30
; Sequence 30, Application US/09309487
; Patent No. 6335318
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/309,487
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-09-309-487-30

Query Match      82.7%; Score 109.2; DB 4; Length 132;
Best Local Similarity 90.0%; Pred. No. 5.2e-31;
Matches 117; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGGAAATGTCCGGACTTTTGGCGGAGAA 60
   |||||
DB 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGGAAATGTCCGGACTTTTGGCGGAGAA 60
   |||||

QY 61 GCGAAGAGTCTAGTCTCGTCTCCGAACTCCAGGAAACGTCTGCTCCTCAAAACGGTTGAC 120
   |||||
DB 61 GCGAAGAGTCTAGTCTCGTCTCCGAACTCCAGGAAACGTCTGCTCCTCAAAACGGTTGAC 120
   |||||

QY 121 AATATTCCG 130
   |||||
DB 121 AATATTCCG 130
   |||||

RESULT 5
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R. M.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```

```
Db 121 AATATTACAG 130

RESULT 4
US-09-967-808-30
; Sequence 30, Application US/09967808
; Patent No. 6514727
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/967,808
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-09-967-808-30

Query Match      82.7%; Score 109.2; DB 4; Length 132;
Best Local Similarity 90.0%; Pred. No. 5.2e-31;
Matches 117; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGGAAATGTCCGGACTTTTGGCGGAGAA 60
   |||||
DB 1 GGACCTTGTCTACTAGTCCCTTACCGAGTAAGGAAATGTACCGGACTTTTGGCGGAGAA 60
   |||||

QY 61 GCGAAGAGTCTAGTCTCGTCTCCGAACTCCAGGAAACGTCTGCTCCTCAAAACGGTTGAC 120
   |||||
DB 61 GGTGAAGTCTCAGTCTCGTCTTCCGAACTCCAGTAAACGTGTGCTCTTAAGACGGCAAC 120
   |||||

QY 121 AATATTCCG 130
   |||||
DB 121 AATATTACAG 130
   |||||

RESULT 5
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R. M.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```

Query Match 22.1%; Score 29.2; DB 3; Length 4403765;
Best Local Similarity 57.8%; Pred. No. 8.5;
Matches 52; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 5 CTCGCTACTAGTCCCTTACCGAGTAAGGAAATGTCCGGACTTTTGGCGGAGAGGCG 64
DB 2191948 CTCGATGACAGCTAGCAAACTAGCGAGATCCGAGCCGGATTCTGACGACGATTCTG 2191889
QY 65 AAAGTCTCAGTCCGCTTCGGAATCCACGG 94
DB 2191888 AACGGCCCGTGCTCAACGAAACAGACGG 2191859

RESULT 6

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 22.1%; Score 29.2; DB 3; Length 4411529;
Best Local Similarity 57.8%; Pred. No. 8.5;
Matches 52; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 5 CTCGCTACTAGTCCCTTACCGAGTAAGGAAATGTCCGGACTTTTGGCGGAGAGGCG 64
DB 2194649 CTCGATGACAGCTAGCAAACTAGCGAGATCCGAGCCGGATTCTGACGACGATTCTG 2194590
QY 65 AAAGTCTCAGTCCGCTTCGGAATCCACGG 94
DB 2194589 AACGGCCCGTGCTCAACGAAACAGACGG 2194560

RESULT 7

US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Query Match 21.5%; Score 28.4; DB 3; Length 4403765;
Best Local Similarity 54.9%; Pred. No. 16;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 2 GACTCGTCTACTAGTCCCTTACCGAGTAAGGAAATGTCCGGACTTTTCCGGCGGAGAG 61
DB 1245607 GACTTCGCCAACAGGCGCTGTTCGGCGATCGCGAAGCAGTTCGGCGACACGCCGAGAA 1245666
QY 62 GCGAAAGTCTCAGTCGCTTCCGAAATCCACGGAAACGTCGTG 103
DB 1245667 CCGCTAACTAAGTCGCTGACGAAGTCCACCACGACGTCGG 1245708

RESULT 8

US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 21.5%; Score 28.4; DB 3; Length 4411529;
Best Local Similarity 54.9%; Pred. No. 16;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 2 GACTCGTCTACTAGTCCCTTACCGAGTAAGGAAATGTCCGGACTTTTCCGGCGGAGAG 61
DB 1246078 GACTTCGCCAACAGGCGCTGTTCGGCGATCGCGAAGCAGTTCGGCGACACGCCGAGAA 1246137
QY 62 GCGAAAGTCTCAGTCGCTTCCGAACTCCACGGAAACGTCGTG 103
DB 1246138 CCGCTAACTAAGTCGCTGACGAGTCCACCACGACGTCGG 1246179

RESULT 9

US-09-306-828-34
; Sequence 34, Application US/09306828
; Patent No. 6475724
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; APPLICANT: Chen, Fu
; APPLICANT: Chen, Hua
; TITLE OF INVENTION: Nucleic Acids, Kits, And Methods For The Diagnosis, Prognosis And
; CURRENT APPLICATION NUMBER: US/09/306,828
; CURRENT FILING DATE: 1999-05-07
; EARLIER APPLICATION NUMBER: US 09/227,881
; EARLIER FILING DATE: 1999-01-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 34
; LENGTH: 5271
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-306-828-34

```
Query Match      20.9%; Score 27.6; DB 4; Length 5271;
Best Local Similarity 56.7%; Pred. No. 2.6;
Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 21 TTACCGAGTAAGGAATGTCGGGACTTTTGGCGGAGAGGCGAAAGTCTCAGTCGCTC 80
Db 2040 TTATTGAGTACTTATATCTGCCAGACACAGAGACAAATGTTGAGCAAGCAGTCAGTCG 2099

QY 81 TCCGAACCTCCACGAAACGTCCTCCTCA 110
Db 2100 CCTACCTTCGTGGAGGTGACAGTTTCTCA 2129

RESULT 10
US-08-938-669A-1
; Sequence 1, Application US/08938669A
; Patent No. 6171788
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND TREATMENT OF GLAUCOMA AND
; TITLE OF INVENTION: RELATED DISEASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2402
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,669A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/791,154
; FILING DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mendelson, Elliot
; REGISTRATION NUMBER: P-42,878
; REFERENCE/DOCKET NUMBER: 07425-0034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 383-6857
; TELEFAX: 202 383-6610
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5300 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-938-669A-1

Query Match      20.9%; Score 27.6; DB 3; Length 5300;
Best Local Similarity 56.7%; Pred. No. 2.6;
Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 21 TTACCGAGTAAGGAATGTCGGGACTTTTGGCGGAGAGGCGAAAGTCTCAGTCGCTC 80
Db 2040 TTATTGAGTACTTATATCTGCCAGACACAGAGACAAATGTTGAGCAAGCAGTCAGTCG 2099

QY 81 TCCGAACCTCCACGAAACGTCCTCCTCA 110
Db 2100 CCTACCTTCGTGGAGGTGACAGTTTCTCA 2129

RESULT 11
```

```
US-09-306-828-1
; Sequence 1, Application US/09306828
; Patent No. 6475724
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; APPLICANT: Chen, Pu
; APPLICANT: Chen, Hua
; TITLE OF INVENTION: Nucleic Acids, Kits, And Methods For The Diagnosis, Prognosis And
; CURRENT APPLICATION NUMBER: US/09/306,828
; CURRENT FILING DATE: 1999-05-07
; EARLIER FILING DATE: 1999-01-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 1
; LENGTH: 5300
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-306-828-1

Query Match      20.9%; Score 27.6; DB 4; Length 5300;
Best Local Similarity 56.7%; Pred. No. 2.6;
Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 21 TTACCGAGTAAGGAATGTCGGGACTTTTGGCGGAGAGGCGAAAGTCTCAGTCGCTC 80
Db 2040 TTATTGAGTACTTATATCTGCCAGACACAGAGACAAATGTTGAGCAAGCAGTCAGTCG 2099

QY 81 TCCGAACCTCCACGAAACGTCCTCCTCA 110
Db 2100 CCTACCTTCGTGGAGGTGACAGTTTCTCA 2129

RESULT 12
US-08-938-669A-2
; Sequence 2, Application US/08938669A
; Patent No. 6171788
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND TREATMENT OF GLAUCOMA AND
; TITLE OF INVENTION: RELATED DISEASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2402
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,669A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/791,154
; FILING DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mendelson, Elliot
; REGISTRATION NUMBER: P-42,878
; REFERENCE/DOCKET NUMBER: 07425-0034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 383-6857
; TELEFAX: 202 383-6610
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
```



```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5304 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-938-669A-2

Query Match      20.9%; Score 27.6; DB 3; Length 5304;
Best Local Similarity 56.7%; Pred. No. 2.6;
Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 21 TTACCGAGTAAGGAAATGTCGGGACTTTTGGCGGAGAGCGGAAAGTCTCAGTCGCTC 80
Db 2040 TTATTGAGTACTTATATCTGCCAGACACGAGACAAAATGGTGAGCAAGCAGTCACTG 2099

QY 81 TCCGAACCTCCACGAAACGTCGTCTCTCA 110
Db 2100 CCTACTCTCGTGAGGTGACAGTTTCTCA 2129

RESULT 13
US-09-306-828-2
; Sequence 2, Application US/09306828
; Patent No. 6475724
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; APPLICANT: Chen, Pu
; APPLICANT: Chen, Hua
; TITLE OF INVENTION: Nucleic Acids, Kits, And Methods For The Diagnosis, Prognosis And
; CURRENT APPLICATION NUMBER: US/09/306,828
; EARLIER FILING DATE: 1999-05-07
; EARLIER FILING DATE: 1999-01-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 2
; LENGTH: 5304
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-306-828-2

Query Match      20.9%; Score 27.6; DB 4; Length 5304;
Best Local Similarity 56.7%; Pred. No. 2.6;
Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 21 TTACCGAGTAAGGAAATGTCGGGACTTTTGGCGGAGAGCGGAAAGTCTCAGTCGCTC 80
Db 2040 TTATTGAGTACTTATATCTGCCAGACACGAGACAAAATGGTGAGCAAGCAGTCACTG 2099

QY 81 TCCGAACCTCCACGAAACGTCGTCTCTCA 110
Db 2100 CCTACTCTCGTGAGGTGACAGTTTCTCA 2129

RESULT 14
US-08-938-669A-3
; Sequence 3, Application US/08938669A
; Patent No. 6171788
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND TREATMENT OF GLAUCOMA AND
; TITLE OF INVENTION: RELATED DISEASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2402

; SEQUENCE CHARACTERISTICS:
; LENGTH: 5304 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-938-669A-2

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ For Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,669A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/791,154
FILING DATE: 28-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mendelson, Elliot
REGISTRATION NUMBER: P-42,878
REFERENCE/DOCKET NUMBER: 07425-0034
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 383-6857
TELEFAX: 202 383-6610
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6169 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-938-669A-3

Query Match      20.9%; Score 27.6; DB 3; Length 6169;
Best Local Similarity 56.7%; Pred. No. 2.8;
Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 21 TTACCGAGTAAGGAAATGTCGGGACTTTTGGCGGAGAGCGGAAAGTCTCAGTCGCTC 80
Db 2040 TTATTGAGTACTTATATCTGCCAGACACGAGACAAAATGGTGAGCAAGCAGTCACTG 2099

QY 81 TCCGAACCTCCACGAAACGTCGTCTCTCA 110
Db 2100 CCTACTCTCGTGAGGTGACAGTTTCTCA 2129

RESULT 15
US-09-306-828-3
; Sequence 3, Application US/09306828
; Patent No. 6475724
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; APPLICANT: Chen, Pu
; APPLICANT: Chen, Hua
; TITLE OF INVENTION: Nucleic Acids, Kits, And Methods For The Diagnosis, Prognosis And
; CURRENT APPLICATION NUMBER: US/09/306,828
; CURRENT FILING DATE: 1999-05-07
; EARLIER FILING DATE: 1999-01-11
; EARLIER FILING DATE: 1999-01-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 3
; LENGTH: 6169
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-306-828-3

Query Match      20.9%; Score 27.6; DB 4; Length 6169;
Best Local Similarity 56.7%; Pred. No. 2.8;
Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 21 TTACCGAGTAAGGAAATGTCGGGACTTTTGGCGGAGAGCGGAAAGTCTCAGTCGCTC 80
Db 2040 TTATTGAGTACTTATATCTGCCAGACACGAGACAAAATGGTGAGCAAGCAGTCACTG 2099

QY 81 TCCGAACCTCCACGAAACGTCGTCTCTCA 110
Db 2100 CCTACTCTCGTGAGGTGACAGTTTCTCA 110
```

Tue Oct 12 09:20:50 2004

us-10-009-317a-31.rni

Page 6

Db 2100 CCCTACCTTCGTCGAGGTGACAGTTCTCA 2129

Search completed: October 11, 2004, 10:56:45
Job time : 51.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 11, 2004, 10:08:26 ; Search time 214 Seconds
(without alignments)
3127.023 Million cell updates/sec

Title: US-10-009-317A-31
Perfect score: 132
Sequence: 1 ggactctgtctactagtcccc.....cggtgacaattattccgca 132

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

ALIGNMENTS

RESULT 1
US-10-313-994-31 Application US/10313994
; Publication No. US20030162718A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/10/313,994
; CURRENT FILING DATE: 2002-12-05
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-10-313-994-31

Query Match 100.0%; Score 132; DB 15; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.7e-40;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTCGCGACTTTTCCGCGAGAA 60
Db 1 GGACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTCGCGACTTTTCCGCGAGAA 60
Qy 61 GGCGAAAGTCTCAGTCCCTCTCCGAACGCCAGAAACGTCTCTCTCAACGGTTGAC 120
Db 61 GGCGAAAGTCTCAGTCCCTCTCCGAACGCCAGAAACGTCTCTCTCAACGGTTGAC 120

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|----|---------------------|
| 1 | 132 | 100.0 | 132 | 15 | US-10-313-994-31 |
| 2 | 109.2 | 82.7 | 132 | 15 | US-10-313-994-30 |
| 3 | 31 | 23.5 | 1197 | 13 | US-10-225-066A-479 |
| 4 | 31 | 23.5 | 1197 | 16 | US-10-225-067-37 |
| 5 | 31 | 23.5 | 1197 | 16 | US-10-374-780A-443 |
| 6 | 29.2 | 22.1 | 82933 | 15 | US-10-080-170-645 |
| 7 | 29.2 | 22.1 | 82933 | 17 | US-10-080-170-645 |
| 8 | 29 | 22.0 | 5493 | 16 | US-10-062-674-2175 |
| 9 | 28.8 | 21.8 | 1739 | 16 | US-10-369-493-36378 |
| 10 | 28.8 | 21.8 | 1983 | 9 | US-09-738-626-2904 |
| 11 | 28.8 | 21.8 | 3309400 | 9 | US-09-738-626-1 |
| 12 | 28.4 | 21.5 | 846 | 15 | US-10-198-846-9562 |
| 13 | 28.4 | 21.5 | 2301 | 16 | US-10-369-493-33481 |
| 14 | 28 | 21.2 | 2574 | 16 | US-10-369-493-34295 |

QY 121 AATATTTCGCA 132
|||||
Db 121 AATATTTCGCA 132

RESULT 2

US-10-313-994-30
; Sequence 30, Application US/10313994
; Publication No. US20030162718A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/10/313,994
; CURRENT FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
US-10-313-994-30

Query Match 82.7%; Score 109.2; DB 15; Length 132;
Best Local Similarity 90.0%; Pred. No. 1e-31;
Matches 117; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 GGACCTGCTACTAGTCCTTCCGAGTAAGGAATGTCGGGAGAA 60
|||||
Db 1 GGACCTGCTACTAGTCCTTCCGAGTAAGGAATGTCGGGAGAA 60
|||||
QY 61 GCGAAAGTCTAGTCGCTCCGAACTCCAGGAAAGCTGTCCTCAAAACGGTTGAC 120
|||||
Db 61 GGTGAAGTCTAGTCGCTTCCGAACTCCAGGAAAGCTGTCCTCAAAACGGTTGAC 120
|||||
QY 121 AATATTTCG 130
|||||
Db 121 AATATTACAG 130

RESULT 3

US-10-225-066A-479
; Sequence 479, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 479
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-479

Query Match 23.5%; Score 31; DB 13; Length 1197;
Best Local Similarity 56.3%; Pred. No. 0.24;
Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
QY 21 TTACCGAGTAAGGAATGTCGGGACTTTTGGCGGAGAGCGCAAGTCTAGTCGTC 80
|||||
Db 987 TTCACAAACATGACATGTTTCTGCGATATTCGAGTAGGAGTCGCTATTAGTTCATC 1046
|||||
QY 81 TCGAACTCCACGGAACGTCCTCTCTCAACGGTTGACAA 123
|||||
Db 1047 TAAGCATTCATGACCGTTTGGTAGCAAGCGTTTGAGAA 1089
|||||

RESULT 4

US-10-225-067-37
; Sequence 37, Application US/10225067
; Publication No. US20040019925A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Creelman, Robert A.
; APPLICANT: Keddle, James
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Dubell, Arnold T.
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Broun, Pierre E
; TITLE OF INVENTION: BIOCHEMISTRY-RELATED POLYNUCLEOTIDES AND
; FILE REFERENCE: POLYPEPTIDES IN PLANTS
; FILE REFERENCE: 51442002042
; CURRENT APPLICATION NUMBER: US/10/225,067
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(1197)
US-10-225-067-37

Query Match 23.5%; Score 31; DB 16; Length 1197;
Best Local Similarity 56.3%; Pred. No. 0.24;
Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 21 TTACCGAGTAAGGAATGTCGGGACTTTTGGCGGAGAAAGCGAAAGTCTCAGTCGCTC 80
Db 987 TTCACAAACATGACATGTTTCTGCGATATCTCGATAGAGTCGCTATTAGTTTATC 1046
QY 81 TCGAACTCCAGGAAAGTCTGCTCCTCAAGCGTTGCAAT 123
Db 1047 TAAGCATTCGAATGAACGGTTTGGTCAGCAAGCGTTTGAGAAAT 1089

RESULT 5
US-10-374-780A-443
; Sequence 443, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Xu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 443
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G2520
US-10-374-780A-443

Query Match 23.5%; Score 31; DB 16; Length 1197;
Best Local Similarity 56.3%; Pred. No. 0.24; Indels 45; Mismatches 0; Gaps 0;
Matches 58; Conservative 0;
QY 21 TTACCGAGTAAGGAATGTCGGGACTTTTGGCGGAGAAAGCGAAAGTCTCAGTCGCTC 80
Db 987 TTCACAAACATGACATGTTTCTGCGATATCTCGATAGAGTCGCTATTAGTTTATC 1046
QY 81 TCGAACTCCAGGAAAGTCTGCTCCTCAAGCGTTGCAAT 123
Db 1047 TAAGCATTCGAATGAACGGTTTGGTCAGCAAGCGTTTGAGAAAT 1089

RESULT 6
US-10-080-170-645/c
; Sequence 645, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 645
; LENGTH: 82993
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-645
Query Match 22.1%; Score 29.2; DB 15; Length 82993;
Best Local Similarity 57.8%; Pred. No. 3.9; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 38;
QY 5 CTGCTCTACTAGTCCCTTACCGAGTAAGGAAATGTGCGGACTTTTCCGGCGAGAAAGCG 64
Db 79038 CTGATGACACGTCAGCAAACTAGCGAGATCCGAGCGGAGTTCTGACGACGCATTCG 78979
QY 65 AAAGTCTCAGTCGCTTCCGAATCCACGG 94
Db 78978 AACGGCCCCGTGGCTCAACGACGACGCG 78949

RESULT 7
US-10-080-170-645/c
; Sequence 645, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 645
; LENGTH: 82993
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-645

Query Match 22.1%; Score 29.2; DB 17; Length 82993;
Best Local Similarity 57.8%; Pred. No. 3.9; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 38;
QY 5 CTGCTCTACTAGTCCCTTACCGAGTAAGGAAATGTGCGGACTTTTCCGGCGAGAAAGCG 64
Db 79038 CTGATGACACGTCAGCAAACTAGCGAGATCCGAGCGGAGTTCTGACGACGCATTCG 78979
QY 65 AAAGTCTCAGTCGCTTCCGAATCCACGG 94
Db 78978 AACGGCCCCGTGGCTCAACGACGACGCG 78949

RESULT 8


```
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match      21.8%; Score 28.8; DB 9; Length 3309400;
Best Local Similarity 58.0%; Pred. No. 16;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 25 CGAGTAAGGAATGTCCGGGACTTTTGGCGGAGAGGGAAGTCTCAGTCGCTCTCCG 84
DB 2811789 CGCGAACGCAATCTCGCGACATTTGGCGGAGAGCAACATTTGTCGGCGCGCACACTG 2811730

QY 85 AACTCCAGCAAGAACTGTCTCTCTCAAA 112
DB 2811729 TCGAACGGTTTACGCTCTCTCAAA 2811702

RESULT 12
US-10-198-846-9562/c
; Sequence 9562, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9562
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 3, 562, 567, 597, 615, 621, 651, 656, 660, 667, 679, 682,
; LOCATION: 690, 707, 723, 724, 753, 762, 790, 794, 795, 796, 799, 800,
; LOCATION: 811, 819, 835, 866, 878, 879, 882
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9562

Query Match      21.5%; Score 28.4; DB 15; Length 882;
Best Local Similarity 60.3%; Pred. No. 2.2;
Matches 47; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 51 GCGGCGAGAGGGAAGTCTCAGTCGCTCTCGAACTCCACGGAAGCTGTGCTCTCA 110
DB 513 GCGGCGAGGAGGAACCGGCTAGATACCTTCCACCCGTGGAACTTTTGTTCTTA 454

QY 111 AACGGTTGACATATTC 128
DB 453 CCTCTTTGACGAAAATC 436

RESULT 13
US-10-369-493-33481
; Sequence 33481, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
```

```
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 33481
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Desulfitobacterium hafniense
US-10-369-493-33481

Query Match      21.5%; Score 28.4; DB 16; Length 2301;
Best Local Similarity 60.3%; Pred. No. 2.9;
Matches 47; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 21 TTACCGAGTAAGGAATGTCCGGGACTTTTGGCGGAGAGGCGAAAGTCTCAGTCGCTC 80
DB 2108 TCACTCTGAGAGGCAATGTTCTGCCCATTTGGCGGAGTGAAGGAAAAGTTTCTCGCGCTC 2167

QY 81 TCCGAACCTCCACGGAAC 98
DB 2168 ACCGGCGGGAATTAAAC 2185

RESULT 14
US-10-369-493-34295
; Sequence 34295, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 34295
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Sphingomonas aromaticivorans
US-10-369-493-34295

Query Match      21.2%; Score 28; DB 16; Length 2574;
Best Local Similarity 52.6%; Pred. No. 4.2;
Matches 61; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 3 ACCTCGCTACTAGTCCTCTTACCGAGTAAGGAATGTCCGGGACTTTTGGCGGGAAGG 62
DB 2159 ACCTCGGAGGCACTCTCGCAATCTGGAGGAGGCGGAGGAGTGCAGAGCGTCGAGC 2218

QY 63 CGAAAGTCTCAGTCGCTCTCGAACTCCACGGAACGTCCTGCTCTCTCAAGCGTTG 118
DB 2219 CGCAAGTCATGGAGCTGCTGCGGCTCATTTCCGGCCCGGAGTTCTCTCAACCGTCTG 2274

RESULT 15
US-10-424-599-66646
; Sequence 66646, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
```

; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 66646
; LENGTH: 221
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_31195C.1
US-10-424-599-66646

Query Match 21.1%; Score 27.8; DB 13; Length 221;
Best Local Similarity 57.5%; Pred. No. 2.5;
Matches 50; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
Qy 40 GCCGGACTTTTGGCGGAGAGCGAAAGTCTCAGTCGCTCTCCGAACTCCACGGAACG 99
Db 26 GCTGGAGTTTGTAGCCGCGGAGGTGACCTTGGGATGGCCCTTCCCAATTCAGTGCTTATT 85
Qy 100 TCTGCTCCTCAACGGTTGACAATATT 126
Db 86 CCTGCTCTTCAGACTGATGGAATATT 112

Search completed: October 11, 2004, 12:14:35
Job time : 217 secs

Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=RCL-HT0881-130
 900-014-b1&t3=2000-09-13&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 28
 High quality sequence stop: 192.

FEATURES

Location/Qualifiers
 1..193
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HT0881"

/note="Organ: head neck; Vector: puc18; Site: 1: SmaI;
 Site: 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 24.4%; Score 32.2; DB 10; Length 193;
 Best Local Similarity 56.0%; Pred. No. 11;
 Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
 Oy 3 ACTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTCGGGACTTTTCGGCGGAGAGG 62
 Db 111 ACCTCAACTCCTGGCTCAATAGTGAAGTAAATCTGTTTCAACAATCAAGAAGACTG 52
 Oy 63 CGAAAGTCTCAGTCGTCTCCGAACTCCAGGAAAGTCTGCTCCCTCAA 111
 Db 51 TGCAGTCATGTCGTCTGATTCATCTGAAACGAGGATCCTCTA 3

RESULT 2
 AA757359/c
 LOCUS
 DEFINITION
 ab96d12.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
 IMAGE:1326935.3. Similar to gb:X59357.60S RIBOSOMAL PROTEIN L22
 (HUMAN); mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 397)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps@mail.nih.gov
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 478 Std Error: 0.00
 Seq primer: -40m13 fwd. RT from Amersham
 High quality sequence stop: 391.

FEATURES

Location/Qualifiers
 1..397
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1326935"
 /lab_host="DH10B"
 /clone_lib="Soares_NFL_T_GBC_S1"

/note="Organ: pooled; Vector: pRT3D-Pac (Pharmacia) with
 a modified polylinker; Site: 1: Not I; Site: 2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBH19W, testis NHT, and B-cell
 NCI CGAP GCBI) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."

Query Match 23.0%; Score 30.4; DB 9; Length 397;
 Best Local Similarity 55.8%; Pred. No. 52;
 Matches 58; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 Oy 21 TTACCGAGTAAGAAATGTCGGGACTTTTCGGCGGAGAGGCGAAAGTCTCAGTCGCTC 80
 Db 357 TTTTGGATTCTGGAATTTTTCAGCAATTTCTACGGGAGAGGTTAAAGTCAATGGCAAAA 298
 Oy 81 TCCGAATCCACGGAACGTCCTCTCTCAACGTTGACAATA 124
 Db 297 CTGGAATCTCGGAATGTTGTTTTCACATTCGACGCTTCAGAATA 254

ORIGIN

RESULT 3
 BX437353/c
 LOCUS
 DEFINITION
 BX437353 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP006YO24
 5-PRIME, mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 951)
 Li,W.B., Gruber,C.; Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 3181.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0CAP006BH120P1&cluster=3181.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0CAP006BH120P1.
 Location/Qualifiers
 1..951
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0CAP006YO24"
 /tissue_type="THYMUS"
 /clone_lib="Homo sapiens THYMUS"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

FEATURES

Location/Qualifiers
 1..951
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0CAP006YO24"
 /tissue_type="THYMUS"
 /clone_lib="Homo sapiens THYMUS"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

FEATURES

Location/Qualifiers
 1..951
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0CAP006YO24"
 /tissue_type="THYMUS"
 /clone_lib="Homo sapiens THYMUS"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

FEATURES

Location/Qualifiers
 1..951
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0CAP006YO24"
 /tissue_type="THYMUS"
 /clone_lib="Homo sapiens THYMUS"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN

Query Match 22.7%; Score 30; DB 13; Length 951;
 Best Local Similarity 45.9%; Pred. No. 99;
 Matches 34; Conservative 17; Mismatches 23; Indels 0; Gaps 0;


```

REFERENCE
AUTHORS      1 (bases 1 to 555)
              Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
              Lewis,S. and Rubin,G.M.
TITLE        BDGP/HMI Drosophila EST Project
JOURNAL      Unpublished (2001)
COMMENT      Contact: Stapleton, M.
              BDGP
              Lawrence Berkeley National Lab
              One Cyclotron Rd, Berkeley, CA 94720, USA
              Fax: 510 486 6798
              Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
              hit genomic sequence AC006495; hit P element sequence 1(3)neo48
              Plate: 441 row: D column: 10
              High quality sequence stop: 522.
FEATURES
source       Location/Qualifiers
              1..555
                /organism="Drosophila melanogaster"
                /mol_type="mRNA"
                /db_xref="taxon:7227"
                /clone="LD4146"
                /sex="male and female"
                /dev_stage="0 to 24 hours mixed stage embryonic"
                /lab_host="X11 Blue"
                /clone_lib="LD Drosophila melanogaster embryo pOT2"
                /note="Organ: embryo; Vector: pOT2; Site: 1; EcoRI; Site 2:
                XhoI; Sized fractionated cDNAs were directly ligated into
                pOT2. "
ORIGIN
Query Match      22.4%; Score 29.6; DB 9; Length 555;
Best Local Similarity 61.8%; Pred. No. 1.1e+02;
Matches 47; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 GGACCTGCTACTAGTCCCTTACCGAGTAAGGAATGTGCCGACTTTTGGCGCAGAA 60
    |||||
DB 216 GGCCCGGAAAAGTTGGCCCTGTACGTACGATATCTGCTGCACGTTGGCGCCAGAA 275
    |||||
QY 61 GCGCAAGTCTCAGTC 76
    |||||
DB 276 GCGGCACAGACATTC 291
    |||||

RESULT 7
A1512018      578 bp mRNA linear EST 19-APR-2001
LOCUS         LD43627.5prime LD Drosophila melanogaster embryo pOT2 Drosophila
DEFINITION    melanogaster cDNA clone LD43627 5prime, mRNA sequence.
ACCESSION     A1512018
VERSION       A1512018.1 GI:4421436
KEYWORDS
SOURCE        Drosophila melanogaster (fruit fly)
ORGANISM      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
REFERENCE     1 (bases 1 to 578)
AUTHORS       Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
              Lewis,S. and Rubin,G.M.
TITLE        BDGP/HMI Drosophila EST Project
JOURNAL      Unpublished (2001)
COMMENT      Contact: Stapleton, M.
              BDGP
              Lawrence Berkeley National Lab
              One Cyclotron Rd, Berkeley, CA 94720, USA
              Fax: 510 486 6798
              Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
              hit genomic sequence AC006495
              Plate: 436 row: C column: 3
              High quality sequence stop: 446.
FEATURES
source       Location/Qualifiers
              1..578
                /organism="Drosophila melanogaster"
                /mol_type="mRNA"

```

```

/db_xref="taxon:7227"
/clone="LD43627"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="X11 Blue"
/clone_lib="LD Drosophila melanogaster embryo pOT2"
/note="Organ: embryo; Vector: pOT2; Site 1: EcoRI; Site 2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2. "
ORIGIN
Query Match      22.4%; Score 29.6; DB 9; Length 578;
Best Local Similarity 61.8%; Pred. No. 1.1e+02;
Matches 47; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 GGACCTGCTACTAGTCCCTTACCGAGTAAGGAATGTGCCGACTTTTGGCGCAGAA 60
    |||||
DB 38 GGCCCGGAAAAGTTGGCCCTGTACGTACGATATCTGCTGCACGTTGGCGCCAGAA 97
    |||||
QY 61 GCGCAAGTCTCAGTC 76
    |||||
DB 98 GCGGCACAGACATTC 113
    |||||

RESULT 8
A1518256      609 bp mRNA linear EST 19-APR-2001
LOCUS         LD37723.5prime LD Drosophila melanogaster embryo pOT2 Drosophila
DEFINITION    melanogaster cDNA clone LD37723 5prime, mRNA sequence.
ACCESSION     A1518256
VERSION       A1518256.1 GI:4424110
KEYWORDS
SOURCE        Drosophila melanogaster (fruit fly)
ORGANISM      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
REFERENCE     1 (bases 1 to 609)
AUTHORS       Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
              Lewis,S. and Rubin,G.M.
TITLE        BDGP/HMI Drosophila EST Project
JOURNAL      Unpublished (2001)
COMMENT      Contact: Stapleton, M.
              BDGP
              Lawrence Berkeley National Lab
              One Cyclotron Rd, Berkeley, CA 94720, USA
              Fax: 510 486 6798
              Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
              hit genomic sequence AC006495; hit P element sequence 1(3)neo48
              Plate: 377 row: B column: 11
              High quality sequence stop: 500.
FEATURES
source       Location/Qualifiers
              1..609
                /organism="Drosophila melanogaster"
                /mol_type="mRNA"
                /db_xref="taxon:7227"
                /clone="LD37723"
                /sex="male and female"
                /dev_stage="0 to 24 hours mixed stage embryonic"
                /lab_host="X11 Blue"
                /clone_lib="LD Drosophila melanogaster embryo pOT2"
                /note="Organ: embryo; Vector: pOT2; Site 1: EcoRI; Site 2:
                XhoI; Sized fractionated cDNAs were directly ligated into
                pOT2. "
ORIGIN
Query Match      22.4%; Score 29.6; DB 9; Length 609;
Best Local Similarity 61.8%; Pred. No. 1.1e+02;
Matches 47; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 GGACCTGCTACTAGTCCCTTACCGAGTAAGGAATGTGCCGACTTTTGGCGCAGAA 60
    |||||
DB 208 GGCCCGGAAAAGTTGGCCCTGTACGTACGATATCTGCTGCACGTTGGCGCCAGAA 267
    |||||

```

```

QY 61 GCGGAAGTCTCAGTC 76
Db 268 GCGGCACAGACATTC 283

RESULT 9
AA941892
LOCUS LD27222.5prime LD Drosophila melanogaster embryo pot2 Drosophila
DEFINITION melanogaster cDNA clone LD27222 5prime, mRNA sequence.
ACCESSION AA941892
VERSION AA941892.1 GI:3102443
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 750)
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST\_estofruitfly.berkeley.edu
hit P element sequence l(3)neod48-3[001],
Plate: 272 row: B column: 10
High quality sequence stop: 627.
Location/Qualifiers
1..750
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="LD27222"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="X11 Blue"
/clone_lib="LD Drosophila melanogaster embryo pot2"
/notes="Organ: embryo; Vector: pot2; Site 1: EcoRI; Site 2:
XhoI; Sized fractionated cDNAs were directly ligated into
pot2."

ORIGIN
Query Match 22.4%; Score 29.6; DB 9; Length 750;
Best Local Similarity 61.8%; Pred. No. 1.2e+02;
Matches 47; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 GGACCTCGTCTACTGCTTACCGAGTAAGAAATGTGCGGACTTTTGGCGGAGAA 60
Db 514 GGCCCGGAAAAGTTGGCCCTGTACGTGTACGAATATCTGCTGCAGTTGGCGCCAGAA 573

QY 61 GCGGAAGTCTCAGTC 76
Db 574 GCGGCACAGACATTC 589

RESULT 10
CA361075/c
LOCUS CA361075.1 NCCGWA 1RT Oncorhynchus mykiss cDNA clone 1RT3405_B_D03 5',
DEFINITION mRNA sequence.
ACCESSION CA361075
VERSION CA361075.1 GI:24665544
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;

Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 244)
Rexroad,C.E. and Keefe,J.W.
Sequence analysis of a rainbow trout normalized cDNA library
Unpublished (2002)
Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Leetown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 x2129
Fax: 304 725 0351
Email: crexroad@nccwa.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified by
cross match v0.990329.
Seq primer: AGCGATAACATTTTCACACAGGA.
Location/Qualifiers
1..244
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="1RT3405_B_D03"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="NCCGWA 1RT"
/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from brain, gill, liver,
spleen, muscle, and kidney."

ORIGIN
Query Match 22.1%; Score 29.2; DB 14; Length 244;
Best Local Similarity 62.2%; Pred. No. 1e-02;
Matches 46; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 32 GGAAATGTGCGGACITTTGCGGAGAGCGAAGTCTCAGTCGCTCCGAACTCCA 91
Db 161 GGAGAAGCTTCGGTCTCTCGGACGACATCTTCCGCTCTCTGCGACACTGGAGACGA 102

QY 92 CGGAACGCTCTGCT 105
Db 101 GGGGAACGGCTGCT 88

RESULT 11
TA41C05Q
LOCUS TA41C05Q
DEFINITION T. brucei sheared genomic DNA clone 41c05, reverse sequence,
genomic survey sequence.
ACCESSION AL455487
VERSION AL455487.1 GI:11855673
KEYWORDS GSS.
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
1 (bases 1 to 517)
Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
Melville,S.E., Rajadream,M.A. and Barrell,B.G.
Direct Submission
Submitted (10-DEC-2000) Trypanosoma brucei Genome Sequencing
project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk
nh@sanger.ac.uk
Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
to give a tight size distribution (
4 Kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org

```

Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES

source
1. .517
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="41c05"

ORIGIN

Query Match 22.1%; Score 29.2; DB 29; Length 517;
Best Local Similarity 52.9%; Pred. No. 1.4e+02;
Matches 64; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 5 CTGCTCTACTAGTCCCTTACCGAGTAAGAAATGTCGGGACTTTTGGCGGAGAGGGG 64
DB 396 CTCTGTGCTCAAGCTTTTTCAGAACTGAATCTTTTGGACATTTTGGCGGGGATGGG 455
QY 65 AAGTCTCAGTCTGCTCCGAATCCACGGAACGCTGCTCTCAAACGGTTGACAATA 124
DB 456 AAAGGATGGCTTCAACGCTGGAGATTGTGGGAAATATTGATTACCGCATCGTAGACAATA 515
QY 125 TT 126
DB 516 TT 517

RESULT 12

CC194385 1135 bp DNA linear GSS 08-MAY-2003
LOCUS CH261-110A6 Sp6.1 CH261 Gallus gallus genomic clone CH261-110A6,
DEFINITION genomic survey sequence.
CC194385
VERSION CC194385.1 GI:30443770
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

REFERENCE

1 (bases 1 to 1135)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Kremitzki, C., Higinbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGTGACACTATAG
Class: BAC ends

High quality sequence start: 127
High quality sequence stop: 556.
Location/Qualifiers

FEATURES

source
1. .1135
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-110A6"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library. For library and clone
ordering information: <http://www.chori.org/bacpac>

ORIGIN

Query Match 22.1%; Score 29.2; DB 28; Length 1135;
Best Local Similarity 54.7%; Pred. No. 1.9e-02;
Matches 58; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 3 ACCTCGTCTACTAGTCCCTTACCGAGTAAGAAATGTCGGGACTTTTGGCGGAGAGG 62
DB 744 ACCTCTTCAATACITTTCTTACTAATATTATGCGGCTGTGGCTCTGCAACTCGAATC 803
QY 63 CGAAGTCTCAGTCTGCTCTCCGAATCCACGGAACGCTCTGCTCCT 108
DB 804 CTTTCTCCCTTAGCACCCCATACACCTCTCTCAACTTCTACTTCT 849

RESULT 13

N60849 354 bp mRNA linear EST 22-MAY-2000
LOCUS TGESty23h01.r1 TGRH tachyzoite cDNA Toxoplasma gondii cDNA clone
DEFINITION tgy23h01.r1 5' similar to SW:R16C_ORYSA P46294 40S RIBOSOMAL
PROTEIN S16. ; mRNA sequence.
N60849
VERSION N60849.1 GI:1207000
KEYWORDS EST.
SOURCE Toxoplasma gondii
ORGANISM Toxoplasma gondii
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.

REFERENCE

1 (bases 1 to 354)
AUTHORS Hehl, A., Manger, I., Marra, M., Sibley, L.D., Ajioke, J.A.,
Aslett, M.A., Dietrich, N., Dubuque, R., Hillier, L., Kucaba, T.,
Wan, K.L., Waterston, R.H. and Boothroyd, J.
WashU-Merck-Stanford-NIH Toxoplasma EST project
Unpublished (1996)
Contact: Marra M

TITLE

JOURNAL
COMMENT

WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxowatson.wustl.edu

David Sibley at toxowest@borcim.wustl.edu for further information
relating to organism, clone or library availability.
Seq primer: T3
High quality sequence stop: 224.
Location/Qualifiers

FEATURES

source
1. .354
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/strain="RH"
/db_xref="taxon:5811"
/clone="tgy23h01.r1"
/lab_host="XLI-Blue MRF"
/clone_lib="TGRH Tachyzoite cDNA"
/note="Vector: Lambda ZAP; Site 1: EcoRI; Site 2: XhoI;
Toxoplasma RH strain tachyzoites were grown in human
foreskin fibroblast cultures in vitro. The library was
constructed by K.L. Wan, Cambridge University. cDNAs were
synthesized from polyA RNAs by oligo d(T) priming and
directionally cloned into the EcoRI to XhoI sites of the
Lambda ZapII vector using the ZAP-cDNA synthesis kit
(Stratagene). WARNING: the library contains a small
percentage of cDNAs derived from the human host cells."

ORIGIN

Query Match 22.0%; Score 29; DB 14; Length 354;
Best Local Similarity 62.9%; Pred. No. 1.4e+02;
Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 38 GTGCGGACTTTTGGCGGAGAGGCGAAAGTCTCAGTCTCGTCTCGCAACTCCACGAAA 97
DB 159 GTGCAACGTTTCGCGCGCAAGAGATGCGGTGCGCTCTGACTCTGACTCTGAGGCAAG 218
QY 98 CGTCTGCTCC 107
DB 219 GGGCTGCTCC 228

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 11, 2004, 08:24:41 ; Search time 189 Seconds
(without alignment)
2966.995 Million cell updates/sec

Title: US-10-009-317A-30
Perfect score: 132
Sequence: 1 ggacctgtctactagtcc.....cggcaacaataattacagt 132

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| C 1 | 30.6 | 23.2 | 567 | AHL11967 | Ahl11967 Human cDN |
| 2 | 30.6 | 23.2 | 1679 | AAC93348 | Aac93348 Human sec |
| 3 | 30.6 | 23.2 | 1985 | Aav84497 | Aav84497 Human sec |
| 4 | 30.6 | 23.2 | 1985 | ABA83280 | AbA83280 Human sec |
| 5 | 30.6 | 23.2 | 1985 | ACH04781 | ACH04781 Novel hum |
| 6 | 30.6 | 23.2 | 1985 | ACD44591 | ACD44591 Human cDN |
| 7 | 30.6 | 23.2 | 1987 | AB189553 | Ab189553 Human pol |
| 8 | 30.6 | 23.2 | 2510 | AHL14641 | Ahl14641 Human cDN |
| 9 | 30.6 | 23.2 | 2793 | AAL160107 | Aal160107 Human pol |
| 10 | 30.6 | 23.2 | 2867 | ADC37174 | Adc37174 Nuclear f |
| 11 | 30.6 | 23.2 | 22013 | AAR85635 | Aar85635 Human imm |
| 12 | 30.6 | 23.2 | 22013 | ABAL16084 | Abal16084 Human ner |
| 13 | 30.6 | 23.2 | 22026 | AAR85636 | Aar85636 Human imm |
| 14 | 30.6 | 23.2 | 22026 | ABAL16085 | Abal16085 Human ner |
| 15 | 30.6 | 23.2 | 175737 | ABK83571 | Abk83571 Human cDN |
| 16 | 30 | 22.7 | 3164 | ABL09382 | Ab109382 Drosophil |
| 17 | 29.8 | 22.6 | 295 | ADA71844 | Ada71844 Rice gene |
| 18 | 28.6 | 21.7 | 1623 | AAS77341 | Aas77341 DNA encod |
| 19 | 28.4 | 21.5 | 8278 | AB111584 | Ab111584 Drosophil |
| 20 | 26.8 | 20.3 | 1161 | AB112853 | Ab112853 Drosophil |
| 21 | 26.8 | 20.3 | 1193 | AB108341 | Ab108341 Drosophil |
| 22 | 26.8 | 20.3 | 2656 | AB112882 | Ab112882 Drosophil |
| 23 | 26.8 | 20.3 | 3322 | AB108340 | Ab108340 Drosophil |

| | | | | | | |
|------|------|------|--------|---|----------|--------------------|
| 24 | 26.8 | 20.3 | 3409 | 4 | ABL12852 | Ab112852 Drosophil |
| 25 | 26.8 | 20.3 | 3511 | 4 | ABL11314 | Ab111314 Drosophil |
| 26 | 26.8 | 20.3 | 39746 | 4 | ABL13398 | Ab113398 Drosophil |
| C 27 | 26.8 | 20.3 | 240825 | 4 | AAF24497 | Aaf24497 Human PG- |
| C 28 | 26.6 | 20.2 | 1770 | 4 | AAF61373 | Aaf61373 A. thalia |
| C 29 | 26.6 | 20.2 | 1875 | 6 | ABZ12889 | Abz12889 Arabidops |
| C 30 | 26.6 | 20.2 | 1875 | 7 | ADA68560 | Ada68560 Arabidops |
| C 31 | 26.4 | 20.0 | 426 | 5 | ABV34395 | Abv34395 Human pro |
| C 32 | 26.4 | 20.0 | 445 | 7 | ABZ23482 | Abz23482 Nucleotid |
| C 33 | 26.4 | 20.0 | 485 | 5 | ABV13276 | Abv13276 Human pro |
| C 34 | 26.4 | 20.0 | 53178 | 4 | AAS59543 | Aas59543 Propionib |
| 35 | 26.4 | 20.0 | 53178 | 7 | ACF64472 | Acf64472 Propionib |
| C 36 | 26.4 | 20.0 | 240823 | 9 | ADD69391 | Add69391 Human PG- |
| C 37 | 26.4 | 20.0 | 240825 | 6 | ABQ81802 | Abq81802 Human PG- |
| C 38 | 26.2 | 19.8 | 576 | 6 | ABQ22915 | Abq22915 Oligonuel |
| C 39 | 26.2 | 19.8 | 576 | 6 | ABQ22914 | Abq22914 Oligonuel |
| C 40 | 26.2 | 19.8 | 631 | 3 | AAAF5998 | Aaf5998 DNA encod |
| C 41 | 26.2 | 19.8 | 1177 | 4 | AAF87519 | Aaf87519 Human str |
| C 42 | 26.2 | 19.8 | 1316 | 8 | ACC85077 | Acc85077 Human MBC |
| C 43 | 26.2 | 19.8 | 4161 | 6 | ABQ69442 | Abq69442 Listeria |
| 44 | 26.2 | 19.8 | 4185 | 6 | ABQ67871 | Abq67871 Listeria |
| C 45 | 26.2 | 19.8 | 10235 | 9 | ADC01265 | Adc01265 Enterohae |

ALIGNMENTS

RESULT 1

AAH11967/C
ID AAH11967 standard; cDNA; 567 BP.

XX AC AAH11967;

DT 26-JUN-2001 (first entry)

DE Human cDNA clone (3'-primer) SEQ ID NO:8802.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

XX EP1074617-A2.

PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-00116126.

XX PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 98JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

(HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or

PT diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 3; SEQ ID NO 8802; 2537pp + Sequence Listing; English.

CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:

CC (a) an oligo-dT primer and an oligonucleotide complementary to the

CC complementary strand of a polynucleotide which comprises one of the 5602

CC nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

| | |
|-----------------------|--|
| CC | of an oligonucleotide comprising a sequence complementary to the |
| CC | complementary strand of a polynucleotide which comprises a 5'-end |
| CC | sequence and an oligonucleotide comprising a sequence complementary to a |
| CC | polynucleotide which comprises a 3'-end sequence, where the |
| CC | oligonucleotide comprises at least 15 nucleotides and the combination of |
| CC | the 5'-end sequence/3'-end sequence is selected from those defined in the |
| CC | specification. The primer sets can be used in antisense therapy and in |
| CC | gene therapy. The primers are useful for synthesising polynucleotides, |
| CC | particularly full-length cDNAs. The primers are also useful for the |
| CC | detection and/or diagnosis of the abnormality of the proteins encoded by |
| CC | the full-length cDNAs. The primers allow obtaining of the full-length |
| CC | cDNAs easily without any specialised methods. AAH03166 to AAH13628 and |
| CC | AAH13633 to AAH18742 represent human cDNA sequences; AAB92445 to AAB95893 |
| CC | represent human amino acid sequences; and AAH13629 to AAH13632 represent |
| CC | oligonucleotides, all of which are used in the exemplification of the |
| CC | present invention |
| XX | |
| SQ | Sequence 567 BP; 172 A; 97 C; 177 G; 118 T; 0 U; 3 Other; |
| | |
| Query Match | 23.2%; Score 30.6; DB 4; Length 567; |
| Best Local Similarity | 52.8%; Pred.No.0.37; Mismatches 66; Conservative 0; Indels 0; Gaps 0; |
| | |
| QY | 4 CTTTGCTACTAGTGCCTTACCGAGTAGGAATGTACCGGACTTTGCGGCGAGAAGGT 63 |
| DB | 430 CCTTCACCTCCAGCACCTTGGCCCAACAGGATAAGCTGGATCCCTTGGCCTTCTGAATAT 371 |
| | |
| QY | 64 GAAGTCTCAGTCGCTTTCGGAATCTCAGCTAAGCTGTGCTCCTTAAGACGGCAACAAT 123 |
| DB | 370 CCACGTGTCTTCAGGTTTTCCAGCATTTCCCAAGACCACTTCCTGTGGCTTCCAAAATGCCCTTTATC 311 |
| | |
| QY | 124 ATTAC 128 |
| DB | 310 ATTTT 306 |
| | |
| RESULT 2 | |
| AAC93348 | |
| ID | AAC93348 standard; cDNA; 1679 BP. |
| XX | |
| XC | AAC93348; |
| XX | |
| DT | 16-FEB-2001 (first entry) |
| XX | |
| DE | Human secreted protein cDNA sequence #39. |
| XX | |
| KW | Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral; |
| KW | antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer; |
| KW | vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic; |
| KW | cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder; |
| KW | neurological disease; infection; human; secreted protein; ss. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200058495-A1. |
| XX | |
| PD | 05-OCT-2000. |
| XX | |
| PF | 23-MAR-2000; 2000WO-US007661. |
| XX | |
| PR | 26-MAR-1999; 93US-0126504P. |
| PR | 07-JAN-2000; 2000US-0174847P. |
| XX | |
| PA | (HUMA-) HUMAN GENOME SCI INC. |
| XX | |
| PI | Rosen CA, Ruben SM, Komatsoulis G; |
| XX | |
| DR | WPI: 2000-611720/58. |
| DR | P-PADB; AAB51417. |
| XX | |
| PT | New nucleic acid molecules encoding 45 human secreted proteins for |
| PT | diagnosing, preventing, treating or ameliorating medical conditions and |
| PT | used as food additives or preservatives. |
| XX | |

PI Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C;
 PI Florence K, Lafleur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR;
 PI Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA;
 PI Carter KC;
 XX
 DR WPI; 1999-059865/05.
 DR P-PSDB; AAW88620, AAW88841, AAW88842, AAW88843, AAW88844, AAW88845,
 DR AAW88846.
 XX
 FT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders.
 XX
 PS Claim 4; Page 347-348, 772pp; English.
 XX
 CC The invention relates to nucleic acid sequences (AAV84411 to AAV84633)
 CC encoding human secreted proteins (AAW88534 to AAW88756). The secreted
 CC protein gene sequences are deposited with the ATCC under deposit numbers
 CC ATCC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
 CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
 CC cells comprising recombinant vectors containing the nucleic acid
 CC sequences are used for the recombinant production of the secreted
 CC proteins. The polynucleotide and amino acid sequences are useful for are
 CC useful for preventing, treating or ameliorating medical conditions e.g.
 CC by protein or gene therapy. Pathological conditions can be also diagnosed
 CC by determining the amount of the new polypeptides in a sample or by
 CC determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the polynucleotides, based on
 CC which tissues they are most highly expressed in, and include developing
 CC products for the diagnosis or treatment of cancer, neurodegenerative
 CC disorders, developmental abnormalities and foetal deficiencies, blood
 CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
 CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
 CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia, AIDS,
 CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
 CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung
 CC or thymus, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners. The
 CC present sequence represents a gene encoding a human secreted protein (see
 CC descriptor line for gene number and clone identification)
 XX
 SQ Sequence 1985 BP; 467 A; 542 C; 466 G; 506 T; 0 U; 4 Other;
 Query Match 23.2%; Score 30.6; DB 2; Length 1985;
 Best Local Similarity 52.8%; Pred. No. 0.6;
 Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
 Qy 4 CCTTGTCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAGGT 63
 Db 1523 CCTTCACTCCAGCACCTTCCCAAGACACACTTCCCTTGGCTTCCAAATGGCTTATC 1582
 Qy 64 GAAAGTCTAGTCTGCTTCCGAACTCCACGTAACGCTGCTCCTAAGACGCAACAAT 123
 Db 1593 CCAGTGTCTTACGTTTCCCAAGACACACTTCCCTTGGCTTCCAAATGGCTTATC 1642
 Qy 124 ATTAC 128
 Db 1643 ATTTC 1647
 RESULT 4
 ABA83280
 ID ABA83280 standard; cDNA; 1985 BP.
 XX
 AC ABA83280;
 XX
 DT 07-FEB-2002 (first entry)
 XX
 DE Human secreted protein gene 87 SEQ ID NO:97.
 XX
 KW Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;
 KW dermatological; immunosuppressive; antiinflammatory; immunostimulant;
 KW cytosstatic; cardiant; vascular; anti-angiogenic; ophthalmological;

PR 06-JUN-1997; 97US-0048880P.
 PR 06-JUN-1997; 97US-0048881P.
 PR 06-JUN-1997; 97US-0048882P.
 PR 06-JUN-1997; 97US-0048883P.
 PR 06-JUN-1997; 97US-0048884P.
 PR 06-JUN-1997; 97US-0048885P.
 PR 06-JUN-1997; 97US-0048889P.
 PR 06-JUN-1997; 97US-0048890P.
 PR 06-JUN-1997; 97US-0048901P.
 PR 06-JUN-1997; 97US-0048915P.
 PR 06-JUN-1997; 97US-0048916P.
 PR 06-JUN-1997; 97US-0048917P.
 PR 06-JUN-1997; 97US-0048949P.
 PR 06-JUN-1997; 97US-0048962P.
 PR 06-JUN-1997; 97US-0048963P.
 PR 06-JUN-1997; 97US-0048964P.
 PR 06-JUN-1997; 97US-0048970P.
 PR 06-JUN-1997; 97US-0048971P.
 PR 06-JUN-1997; 97US-0048972P.
 PR 06-JUN-1997; 97US-0048974P.
 PR 06-JUN-1997; 97US-0049019P.
 PR 06-JUN-1997; 97US-0049020P.
 PR 06-JUN-1997; 97US-0049373P.
 PR 06-JUN-1997; 97US-0049374P.
 PR 06-JUN-1997; 97US-0049375P.
 PR 05-SEP-1997; 97US-0057584P.
 PR 05-SEP-1997; 97US-0057627P.
 PR 05-SEP-1997; 97US-0057628P.
 PR 05-SEP-1997; 97US-0057629P.
 PR 05-SEP-1997; 97US-0057634P.
 PR 05-SEP-1997; 97US-0057635P.
 PR 05-SEP-1997; 97US-0057642P.
 PR 05-SEP-1997; 97US-0057643P.
 PR 05-SEP-1997; 97US-0057644P.
 PR 05-SEP-1997; 97US-0057645P.
 PR 05-SEP-1997; 97US-0057646P.
 PR 05-SEP-1997; 97US-0057647P.
 PR 05-SEP-1997; 97US-0057648P.
 PR 05-SEP-1997; 97US-0057649P.
 PR 05-SEP-1997; 97US-0057650P.
 PR 05-SEP-1997; 97US-0057651P.
 PR 05-SEP-1997; 97US-0057654P.
 PR 05-SEP-1997; 97US-0057661P.
 PR 05-SEP-1997; 97US-0057662P.
 PR 05-SEP-1997; 97US-0057663P.
 PR 05-SEP-1997; 97US-0057664P.
 PR 05-SEP-1997; 97US-0057665P.
 PR 05-SEP-1997; 97US-0057666P.
 PR 05-SEP-1997; 97US-0057667P.
 PR 05-SEP-1997; 97US-0057668P.
 PR 05-SEP-1997; 97US-0057760P.
 PR 05-SEP-1997; 97US-0057761P.
 PR 05-SEP-1997; 97US-0057762P.
 PR 05-SEP-1997; 97US-0057763P.
 PR 05-SEP-1997; 97US-0057764P.
 PR 05-SEP-1997; 97US-0057765P.
 PR 05-SEP-1997; 97US-0057766P.
 PR 05-SEP-1997; 97US-0057777P.
 PR 05-SEP-1997; 97US-0057778P.
 PR 18-DEC-1997; 97US-0070923P.
 (HUMA-) HUMAN GENOME SCI INC.
 Young P, Greene JM, Ferrie AX, Ruben SM, Rosen CA, Hu J;

KW neuroprotective; nootropic; anticonvulsant; antialzheimers; vulnery;
KW antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
KW multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
KW human immunodeficiency virus; hyperproliferative disorder; wound healing;
KW Gaucher's disease; cardiovascular disease; Scimitar syndrome; chemotaxis;
KW Chaga's cardiomyopathy; coronary arteriosclerosis; angiogenic disorder;
KW corneal graft neovascularisation; diabetic retinopathy; regeneration;
KW neurological disorder; Huntington's chorea; Alzheimer's disease;
KW Parkinson's disease; infectious disease; ss.
XX Homo sapiens.
XX WO200162891-A2.
XX 30-AUG-2001.
XX 21-FEB-2001; 2001WO-US005614.
XX 24-FEB-2000; 2000US-0184836P.
XX 29-MAR-2000; 2000US-0193170P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ni J, Ebner R, Lafleur DW, Moore PA, Olsen HS, Rosen CA;
PI Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;
PI Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferrie AM, Fan P;
PI Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G, Zeng Z;
PI Greene JW;
XX WPI; 2001-625724/72.
XX P-PSDB; ABB50387.
XX Nucleic acids encoding 207 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating, e.g. cancers, Parkinson's disease
PT and diabetic retinopathy.
XX Claim 1; Page 968-969; 1533pp; English.
XX ABB50301 to ABB51287 and ABB53194 to ABB53441 represent human secreted
CC proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
CC activities based on the tissues and cells the genes are expressed in.
CC Example of these activities include: immunomodulatory; antisclerotic;
CC dermatological; immunosuppressive; anti-inflammatory; immunostimulant;
CC anti-HIV; cytostatic; cardiac; anti-angiogenic; ophthalmological;
CC neuroprotective; nootropic; anticonvulsant; antialzheimers; vascular;
CC antiparkinsonian; antimicrobial; and vulnery. (I) and (II) can be used
CC in gene therapy and vaccine production. (I) and (II) can be used in the
CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
CC Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome,
CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
CC disorders (e.g. corneal graft neovascularisation and diabetic
CC retinopathy), neurological disorders (e.g. Huntington's chorea,
CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. ABB53185 to
CC ABB53193 and ABB50300 represent sequences used in the exemplification of
CC the present invention
XX Sequence 1985 BP; 467 A; 542 C; 466 G; 506 T; 0 U; 4 Other;
XX Query Match 23.2%; Score 30.6; DB 4; Length 1985;
XX Best Local Similarity 52.8%; Pred. No. 0.6;
XX Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
XX 4 CCTTGTCTACTGCTTCCCTACCGAGTAAGGAATGTACCGGACTTTGGCGGAGAGGT 63
XX 1523 CCTTACTCCGACACCTTCCCAACAGATAGCTGGATCCCTTGGCTTCTGAATAT 1582
XX 64 GAAAGTCTCAGTCGCTTTCGAACTCCAGTAAAGTGTGCTCCTTAAGACGCAACAAAT 123
XX 1583 CCCAGTGTCTTCAGGTTTCCCAAGACCACTTCCCTGTGGGCTTCCAAAATGCCCTTTATC 1642

QY 124 ATTAC 128
DB 1643 ATTTC 1647
RESULT 5
ACH04781
ID ACH04781 standard; cDNA; 1985 BP.
XX AC ACH04781;
XX 02-OCT-2003 (first entry)
XX Novel human secreted protein #87 cDNA.
XX Human; ss; gene therapy; autoimmune disorder; multiple sclerosis; cancer;
KW systemic lupus erythematosus; haematopoietic cell disorder; allergy;
KW agammaglobulinaemia; ataxia telangiectasia; blood coagulation disorder;
KW afibrinogenaemia; thrombocytopenia; graft-versus-host disease; arthritis;
KW inflammatory condition; ischaemia-reperfusion injury; infectious disease;
KW hyperproliferative disorder; purpura; viral infection; regeneration;
KW bacterial infection; ulcer; Alzheimer's disease; gene.
XX Homo sapiens.
XX US2003065160-A1.
XX 03-APR-2003.
XX 07-DEC-2001; 2001US-00004860.
XX 06-JUN-1997; 97US-0048875P.
XX 06-JUN-1997; 97US-0048876P.
XX 06-JUN-1997; 97US-0048877P.
XX 06-JUN-1997; 97US-0048878P.
XX 06-JUN-1997; 97US-0048880P.
XX 06-JUN-1997; 97US-0048881P.
XX 06-JUN-1997; 97US-0048882P.
XX 06-JUN-1997; 97US-0048883P.
XX 06-JUN-1997; 97US-0048884P.
XX 06-JUN-1997; 97US-0048885P.
XX 06-JUN-1997; 97US-0048892P.
XX 06-JUN-1997; 97US-0048893P.
XX 06-JUN-1997; 97US-0048894P.
XX 06-JUN-1997; 97US-0048895P.
XX 06-JUN-1997; 97US-0048896P.
XX 06-JUN-1997; 97US-0048897P.
XX 06-JUN-1997; 97US-0048898P.
XX 06-JUN-1997; 97US-0048899P.
XX 06-JUN-1997; 97US-0049000P.
XX 06-JUN-1997; 97US-0048901P.
XX 06-JUN-1997; 97US-0048916P.
XX 06-JUN-1997; 97US-0048917P.
XX 06-JUN-1997; 97US-0048949P.
XX 06-JUN-1997; 97US-0048962P.
XX 06-JUN-1997; 97US-0048963P.
XX 06-JUN-1997; 97US-0048964P.
XX 06-JUN-1997; 97US-0048970P.
XX 06-JUN-1997; 97US-0048971P.
XX 06-JUN-1997; 97US-0048972P.
XX 06-JUN-1997; 97US-0048974P.
XX 06-JUN-1997; 97US-0049019P.
XX 06-JUN-1997; 97US-0049020P.
XX 06-JUN-1997; 97US-0049373P.
XX 06-JUN-1997; 97US-0049374P.
XX 05-SEP-1997; 97US-0057584P.
XX 05-SEP-1997; 97US-0057827P.
XX 05-SEP-1997; 97US-0057828P.
XX 05-SEP-1997; 97US-0057629P.
XX 05-SEP-1997; 97US-0057634P.
XX 05-SEP-1997; 97US-0057635P.

| | | | |
|-----------------------|--|----------------|----------------|
| PR | 05-SEP-1997 | 97US-0057642P. | 97US-0057642P. |
| PR | 05-SEP-1997 | 97US-0057643P. | 97US-0057643P. |
| PR | 05-SEP-1997 | 97US-0057644P. | 97US-0057644P. |
| PR | 05-SEP-1997 | 97US-0057645P. | 97US-0057645P. |
| PR | 05-SEP-1997 | 97US-0057646P. | 97US-0057646P. |
| PR | 05-SEP-1997 | 97US-0057647P. | 97US-0057647P. |
| PR | 05-SEP-1997 | 97US-0057648P. | 97US-0057648P. |
| PR | 05-SEP-1997 | 97US-0057649P. | 97US-0057649P. |
| PR | 05-SEP-1997 | 97US-0057650P. | 97US-0057650P. |
| PR | 05-SEP-1997 | 97US-0057651P. | 97US-0057651P. |
| PR | 05-SEP-1997 | 97US-0057654P. | 97US-0057654P. |
| PR | 05-SEP-1997 | 97US-0057661P. | 97US-0057661P. |
| PR | 05-SEP-1997 | 97US-0057662P. | 97US-0057662P. |
| PR | 05-SEP-1997 | 97US-0057666P. | 97US-0057666P. |
| PR | 05-SEP-1997 | 97US-0057667P. | 97US-0057667P. |
| PR | 05-SEP-1997 | 97US-0057668P. | 97US-0057668P. |
| PR | 05-SEP-1997 | 97US-0057760P. | 97US-0057760P. |
| PR | 05-SEP-1997 | 97US-0057761P. | 97US-0057761P. |
| PR | 05-SEP-1997 | 97US-0057762P. | 97US-0057762P. |
| PR | 05-SEP-1997 | 97US-0057763P. | 97US-0057763P. |
| PR | 05-SEP-1997 | 97US-0057764P. | 97US-0057764P. |
| PR | 05-SEP-1997 | 97US-0057765P. | 97US-0057765P. |
| PR | 05-SEP-1997 | 97US-0057769P. | 97US-0057769P. |
| PR | 05-SEP-1997 | 97US-0057770P. | 97US-0057770P. |
| PR | 05-SEP-1997 | 97US-0057771P. | 97US-0057771P. |
| PR | 05-SEP-1997 | 97US-0057774P. | 97US-0057774P. |
| PR | 05-SEP-1997 | 97US-0057775P. | 97US-0057775P. |
| PR | 05-SEP-1997 | 97US-0057776P. | 97US-0057776P. |
| PR | 05-SEP-1997 | 97US-0057777P. | 97US-0057777P. |
| PR | 05-SEP-1997 | 97US-0057778P. | 97US-0057778P. |
| PR | 18-DEC-1997 | 97US-0070923P. | 97US-0070923P. |
| PR | 04-JUN-1998 | 98MO-US011422. | 98MO-US011422. |
| PR | 15-JUL-1998 | 98US-0092921P. | 98US-0092921P. |
| PR | 30-JUL-1998 | 98US-0094657P. | 98US-0094657P. |
| PR | 04-DEC-1998 | 98US-00205258. | 98US-00205258. |
| XX | | | |
| PA | (HUMA-) HUMAN GENOME SCI INC. | | |
| XX | | | |
| XX | Young P, Greene JM, Ruben SM, Rosen CA, Hu J; | | |
| PI | Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C; | | |
| PI | Florence K, Lafleur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR; | | |
| PI | Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA; | | |
| PI | Carter KC; | | |
| XX | | | |
| DR | WPI; 2003-540804/51. | | |
| DR | P-PSDB; ABO44644. | | |
| XX | | | |
| PT | New isolated protein, useful for preparing a composition for diagnosing | | |
| PT | or treating cancer, inflammatory, immune or infectious diseases. | | |
| XX | | | |
| PS | Example 1; SEQ ID NO 97; 172pp; English. | | |
| XX | | | |
| CC | The invention relates to an isolated HEMA80 protein. The protein is | | |
| CC | useful for preparing a composition for diagnosing or treating autoimmune | | |
| CC | disorders e.g. multiple sclerosis and systemic lupus erythematosus; | | |
| CC | haematopoietic cell disorders e.g. agammaglobulinaemia and staxia | | |
| CC | relangectasia; allergy; graft-versus-host disease; inflammatory | | |
| CC | conditions e.g. ischaemia-reperfusion injury and arthritis; | | |
| CC | hyperproliferative disorders e.g. cancer and purpura; infectious disease | | |
| CC | e.g. viral infection and bacterial infection. The polynucleotide or | | |
| CC | protein can be used to regenerate damaged tissue e.g. ulcers and | | |
| CC | Alzheimer's disease. The present sequence represents a novel human | | |
| CC | secreted protein cDNA. Note: The sequence data for this patent did not | | |
| CC | form part of the printed specification but was obtained in electronic | | |
| CC | format directly from USPTO at | | |
| CC | seqdata.uspto.gov/sequence.html?DocID=20030065160 | | |
| XX | | | |
| SQ | Sequence 1985 BP; 467 A; 542 C; 466 G; 506 T; 0 U; 4 Other; | | |
| Query Match | 23.2%; Score 30.6; DB 8; Length 1985; | | |
| Best Local Similarity | 52.8%; Pred. No. 0.6; | | |
| Matches | 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0; | | |

CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 1987 BP; 464 A; 548 C; 467 G; 507 T; 0 U; 1 Other;

Query Match 23.2%; Score 30.6; DB 6; Length 1987;
 Best Local Similarity 52.8%; Pred. No. 0.6;
 Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CCTTGTCTACTAGTCCCTTACCGAGTAAGGAAATGTACCGGACTTTTTCGGCGGAGAGGT 63
 |||||
 DB 1528 CTTTCACTCCAGACCTTGCACACAGATAGCTGGATCCCTTGGCCTTCGAATAT 1587

QY 64 GAAAGTCTCAGTCGTTTCGGAATCCAGTAAACGTGTGCTCTTAAGACGGCAACAAAT 123
 |||||
 DB 1588 CCCAGTGTCTCAGGTTTCCCAAGACCACTCCCTGTGGGCTTCCAAAATGGCCTTTATC 1647

QY 124 ATTAC 128
 |||||
 DB 1648 ATTTC 1652

RESULT 8
 AAH14641
 ID AAH14641 standard; cDNA; 2510 BP.

XX AC AAH14641;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:12295.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 6; SEQ ID NO 12295; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:

CC (a) an oligo-dT primer and an oligonucleotide complementary to the

CC complementary strand of a polynucleotide which comprises one of the 5602

CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13629 and
 CC AAH13630 to AAH18742 represent human cDNA sequences; AAH92446 to AAH35893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX SQ Sequence 2510 BP; 574 A; 665 C; 630 G; 641 T; 0 U; 0 Other;

Query Match 23.2%; Score 30.6; DB 4; Length 2510;
 Best Local Similarity 52.8%; Pred. No. 0.66;
 Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CCTTGTCTACTAGTCCCTTACCGAGTAAGGAAATGTACCGGACTTTTTCGGCGGAGAGGT 63
 |||||
 DB 2081 CTTTCACTCCAGACCTTGCACACAGATAGCTGGATCCCTTGGCCTTCGAATAT 2140

QY 64 GAAAGTCTCAGTCGTTTCGGAATCCAGTAAACGTGTGCTCTTAAGACGGCAACAAAT 123
 |||||
 DB 2141 CCCAGTGTCTCAGGTTTCCCAAGACCACTCCCTGTGGGCTTCCAAAATGGCCTTTATC 2200

QY 124 ATTAC 128
 |||||
 DB 2201 ATTTC 2205

RESULT 9

AAI60107/c

ID AAI60107 standard; cDNA; 2793 BP.

XX AC AAI60107;

XX 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 4096.

XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
 XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 XX leukaemia; ss.

XX OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

PA (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR P-PSDB; AAM40951.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX Claim 1; SEQ ID NO 4096; 10078pp; English.
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX Sequence 2793 BP; 686 A; 707 C; 743 G; 653 T; 0 U; 4 Other;
SQ Query Match 23.2%; Score 30.6; DB 4; Length 2793;
Best Local Similarity 52.8%; Pred. No. 0.69;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 4 CCTGTGCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAGGT 63
DB 439 CCTTCACTCCAGACCTTCCCAACAGATAGCTGGATCCCTTGGCTTCTGAATAT 380
QY 64 GAAAGTCTCAGTCCGTTCCGAACTCCAGTAAACGTGTCTCTTAAGACGCAACAAT 123
DB 379 CCCAGTGTCTTCAAGTTTCCAGAGCACTTCCCTGTGGGCTTCCAAAATGGCCTTTATC 320
QY 124 ATTAC 128
DB 319 ATTTC 315
RESULT 10
ADC37174
ID ADC37174 standard; DNA; 2867 BP.
XX ADC37174;
XX
XX 18-DEC-2003 (first entry)
XX Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 7.
XX
XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
XX cancer; infectious disease; bone disease; AIDS;
XX neurodegenerative disease; ischaemic disorder; Antinflammatory;
XX Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-Hiv;
XX Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX
XX Homo sapiens.
XX
XX WO2003048202-A2.
XX
XX 12-JUN-2003.
XX
XX 03-DEC-2002; 2002WO-JP012644.

PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX (ASAH) ASAH KASEI KK.
XX Matsuda A, Muramatsu S;
XX WPI; 2003-505282/47.
DR P-PSDB; ADC37175.
XX New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX Claim 4; SEQ ID NO 7; 938pp; English.
XX The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX Sequence 2867 BP; 662 A; 761 C; 747 G; 697 T; 0 U; 0 Other;
SQ Query Match 23.2%; Score 30.6; DB 9; Length 2867;
Best Local Similarity 52.8%; Pred. No. 0.69;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 4 CCTGTGCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAGGT 63
DB 2439 CCTTCACTCCAGACCTTCCCAACAGATAGCTGGATCCCTTGGCTTCTGAATAT 2498
QY 64 GAAAGTCTCAGTCCGTTTCCGAACTCCAGTAAACGTGTCTCTTAAGACGCAACAAT 123
DB 2499 CCCAGTGTCTTCAAGTTTCCAGAGCACTTCCCTGTGGGCTTCCAAAATGGCCTTTATC 2558
QY 124 ATTAC 128
DB 2559 ATTTC 2563
RESULT 11
AAK85635
ID AAK85635 standard; DNA; 22013 BP.
XX AAK85635;
XX
XX 07-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40447.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225211P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233397P.
PR 14-SEP-2000; 2000US-0233398P.
PR 14-SEP-2000; 2000US-0233399P.
PR 14-SEP-2000; 2000US-023400P.
PR 14-SEP-2000; 2000US-023401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234233P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
PS Disclosure; SEQ ID NO 40447; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK6170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For

CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 22013 BP; 5294 A; 5048 C; 5035 G; 6636 T; 0 U; 0 Other;

Query Match 23.2%; Score 30.6; DB 4; Length 22013;
Best Local Similarity 52.8%; Pred. No. 1.5;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 4 CCTTGTCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAAGGT 63
Db 21165 CCTTACTCCAGCACCTTGGCCCAACAGGATAAGCTGGATCCCTTGGCTTCTGAATAT 21224

QY 64 GAAGTCTCAGTGCCTTCCGACTCCAGTAAACGTGTGCTCCTTAGACGGCAACAAT 123
Db 21225 CCAGTGTCTTCAGGTTTCCCAAGACCACTTCCCTGTGGCTTCCAAATGGCTTTATC 21284

QY 124 ATTAC 128
Db 21285 ATTTC 21289

RESULT 12
ID ABA16084
AB ABA16084 standard; DNA; 22013 BP.
AC ABA16084;
XX
XX 23-JAN-2002 (first entry)
XX Human nervous system related polynucleotide SEQ ID NO 8415.
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
XX antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
XX antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
XX antiparasitic; cardiant; immune disorder; cardiovascular disorder;
XX neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
OS Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241185P.
PR 20-OCT-2000; 2000US-0241186P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-024221P.
PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
XX Disclosure; SEQ ID NO 8415; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (AB14678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing

CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 22013 BP; 5294 A; 5048 C; 5035 G; 6636 T; 0 U; 0 Other;
Query Match 23.2%; Score 30.6; DB 5; Length 22013;
Best Local Similarity 52.8%; Pred. No. 1.5;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
Qy 4 CCTTGTCTACTAGTCCCTTACCGAGTAAGGAATGTACCGGACTTTTGGCGGAGAGGT 63
Db 21165 CCTTCACTCCAGCACCTTGGCCCAAGAGTAAGCTGGATCCCTTGGCTTCTGATAT 21224
Qy 64 GAAAGTCTCAGTCGTTCCGAACTCCAGTAAACGTGTGCTCCTTAAGACGGCAACAAT 123
Db 21225 CCCAGTGTCTCAGTTTCCCAAGACCACTTCCCTGTGGGCTTCCAAAATGGCCTTATC 21284
Qy 124 ATTAC 128
Db 21285 ATTTC 21289
RESULT 13
AAK85636
ID AAK85636 standard; DNA; 22026 BP.
AC AAK85636;
XX
XX 07-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40448.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US0001354.
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205151P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 05-SEP-2000; 2000US-0230437P.
PR 05-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231142P.
PR 08-SEP-2000; 2000US-0231143P.
PR 08-SEP-2000; 2000US-0231144P.
PR 08-SEP-2000; 2000US-0231143P.
PR 08-SEP-2000; 2000US-0231144P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231368P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239359P.
PR 13-OCT-2000; 2000US-0239357P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244647P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Disclosure; SEQ ID NO 40448; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
SQ Sequence 22026 BP; 5289 A; 5044 C; 5030 G; 6663 T; 0 U; 0 Other;
Query Match 23.2%; Score 30.6; DB 4; Length 22026;
Best Local Similarity 52.8%; Pred. NO. 1.5;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 4 CCTGTCTACTAGTCCCTTACCAGATAGGAATACCGGACTTTTGGCGGAGAGGT 63

Db 21178 CCTCATCCAGCACCTTCCCAACAGTAAGTGGATCCCTTGGCTTCTGAATAT 21237
Oy 64 GAAAGTCTAGTCGCTTCCGAACCTCACTGCTCTTAAGACGGCAACAAT 123
Db 21238 CCCAGTCTCTCAGGTTTCCCAAGACCACTTCCCTGTGGCTTCCAAAATGGCCTTATC 21297
Oy 124 ATTAC 128
Db 21298 ATTTC 21302

RESULT 14
ABAI6085
ID ABAI6085 standard; DNA; 22026 BP.
XX
AC ABAI6085;
XX
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 8416.
XX
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnary;
KW antiparkinsonian; antiskelting; anianaemic; antiarthritic; cancer;
KW antihemantic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
XX
EN W0200159063-A2.
XX
PD 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251160P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256713P.
 PR 08-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254037P.
 PR 03-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-541565/60.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system cancers
 PT and metastases.
 XX
 XX Disclosure; SEQ ID NO 8416; 1701bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
 CC (AB14678-AB18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 22026 BP; 5289 A; 5044 C; 5030 G; 6663 T; 0 U; 0 Other;
 Query Match 23.2%; Score 30.6; DB 5; Length 22026;
 Best Local Similarity 52.8%; Pred. No. 1.5;
 Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
 QY 4 CCTTGCTACTAGTCCCTTACCGAGTAAAGGAATGTACCGGACTTTTCGGCGAGAGGT 63
 DB |||||
 21178 CTTTACTCCAGACCTTCCCAACAGATAGCTGATCCCTTGGCTTCTGTATAT 21237
 QY 64 GAAAGTCTCAGTGGTTTCCGAACTCCAGTAAAGTGTGTTCTTAAGACGGCAACAT 123
 DB |||||
 21238 CCCAGTGTCTTCAGGTTTCCCAAGACCACTTCCCTGTGGGCTTCCAAATGGCCTTATC 21297
 DB |||||

QY 124 ATTAC 128
 DB |||||
 21298 ATTTC 21302
 RESULT 15
 ABK83571
 ID ABK83571 standard; cDNA; 175737 BP.
 XX
 AC ABK83571;
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Human cDNA differentially expressed in granulocytic cells #142.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX
 OS Homo sapiens.
 XX
 PN WO200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001; 2001WO-US030821.
 XX
 PR 03-OCT-2000; 2000US-0237189P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 XX
 XX WPI; 2002-435328/46.
 XX
 PT Detecting granulocyte activation by detecting differential expression of
 PT genes associated with granulocyte activation, which serves as diagnostic
 PT markers that is useful for monitoring disease states and drug toxicity.
 XX
 PS Claim 1; SEQ ID NO 142; 114bp; English.
 XX
 CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing the
 CC expression level to an expression level in an unactivated GC, where
 CC differential expression of Gs is indicative of GCA. Also included are
 CC modulating (M2) GA by contacting GC with an agent that alters the
 CC expression of at least one gene in Gs; (2) screening (M3) for an agent
 CC capable of modulating GCA or an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease using the gene expression
 CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease, by detecting the level of
 CC expression in a sample of the tissue of gene(s) from Gs, where the level
 CC of expression of the gene is indicative of inflammation; (4) treating
 CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease, by contacting a tissue having inflammation with an
 CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
 CC is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful
 CC for screening an agent capable of modulating GCA preferably in an
 CC inflammation in a tissue; M4 is useful for detecting an inflammation
 CC (especially chronic) in a tissue, an allergic response in a subject,
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's

Tue Oct 12 09:20:48 2004

CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 175737 BP; 41985 A; 43790 C; 42407 G; 47555 T; 0 U; 0 Other;
Query Match 23.2%; Score 30.6; DB 6; Length 175737;
Best Local Similarity 52.8%; Pred. No. 3.4;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 4 CCTTGCTCTACTAGTCCCTTTACCGAGTAAGGAAATGTACCGGACTTTTTCGGCGGAGAAAGGT 63
DB 35519 CCTTCACCTCCAGACACCTTGGCCACAGATAGCTGGATCCCTTGGCCTTCTGAATAT 35578
QY 64 GAAAGTCTCAGTCGCTTTCCGAATCCACGTAACGTTGTCTCTAGACGGCAACAAT 123
DB 35579 CCACGTGCTTCAGGTTTCCCAAGACCACCTTCCCTGTGGGCTTCCAAAATGCCCTTTATC 35638
QY 124 ATTAC 128
DB 35639 ATTTC 35643

Search completed: October 11, 2004, 09:30:38
Job time : 194 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 15:20:02 ; Search time 31.5 Seconds
(without alignments)
37.896 Million cell updates/sec

Title: US-10-009-317A-32

Perfect score: 116

Sequence: 1 GFCRCICTRGFCRCICTR 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------------------|-------------------|
| 1 | 101 | 87.1 | 18 | 3 US-09-309-487-1 | Sequence 1, Appl |
| 2 | 101 | 87.1 | 18 | 4 US-09-967-808-1 | Sequence 1, Appl |
| 3 | 93 | 80.2 | 18 | 4 US-10-141-645-5 | Sequence 5, Appl |
| 4 | 93 | 80.2 | 18 | 4 US-10-141-645-6 | Sequence 6, Appl |
| 5 | 91 | 78.4 | 18 | 3 US-09-309-487-9 | Sequence 9, Appl |
| 6 | 91 | 78.4 | 18 | 4 US-09-967-808-9 | Sequence 9, Appl |
| 7 | 90 | 77.6 | 18 | 4 US-10-141-645-1 | Sequence 1, Appl |
| 8 | 87 | 75.0 | 18 | 4 US-09-917-340-53 | Sequence 53, Appl |
| 9 | 87 | 75.0 | 18 | 4 US-10-141-645-2 | Sequence 2, Appl |
| 10 | 85 | 73.3 | 18 | 4 US-10-141-645-3 | Sequence 3, Appl |
| 11 | 85 | 73.3 | 18 | 4 US-10-141-645-4 | Sequence 4, Appl |
| 12 | 83 | 71.6 | 18 | 4 US-10-141-645-7 | Sequence 7, Appl |
| 13 | 75 | 64.7 | 18 | 4 US-10-141-645-8 | Sequence 8, Appl |
| 14 | 75 | 64.7 | 18 | 4 US-10-141-645-9 | Sequence 9, Appl |
| 15 | 73 | 62.9 | 92 | 3 US-09-309-487-21 | Sequence 21, Appl |
| 16 | 73 | 62.9 | 92 | 4 US-09-967-808-21 | Sequence 21, Appl |
| 17 | 64 | 55.2 | 76 | 3 US-09-309-487-14 | Sequence 14, Appl |
| 18 | 64 | 55.2 | 76 | 3 US-09-309-487-29 | Sequence 29, Appl |
| 19 | 64 | 55.2 | 76 | 4 US-09-967-808-14 | Sequence 14, Appl |
| 20 | 64 | 55.2 | 76 | 4 US-09-967-808-29 | Sequence 29, Appl |
| 21 | 64 | 55.2 | 76 | 4 US-10-141-645-15 | Sequence 15, Appl |
| 22 | 60 | 51.7 | 180 | 4 US-09-510-238A-286 | Sequence 286, App |
| 23 | 58 | 50.0 | 9 | 3 US-09-309-487-18 | Sequence 18, Appl |
| 24 | 58 | 50.0 | 9 | 4 US-09-967-808-18 | Sequence 18, Appl |
| 25 | 58 | 50.0 | 9 | 4 US-10-141-645-36 | Sequence 36, Appl |
| 26 | 56 | 48.3 | 9 | 4 US-10-141-645-49 | Sequence 49, Appl |
| 27 | 56 | 48.3 | 323 | 4 US-09-270-767-41896 | Sequence 41896, A |

| | | | | | | |
|----|------|------|------|---|----------------------|-------------------|
| 28 | 55.5 | 47.8 | 1400 | 3 | US-08-630-915A-37 | Sequence 37, Appl |
| 29 | 55.5 | 47.8 | 1400 | 4 | US-09-879-957-37 | Sequence 37, Appl |
| 30 | 55 | 47.4 | 168 | 4 | US-09-252-991A-32502 | Sequence 32502, A |
| 31 | 55 | 47.4 | 585 | 4 | US-09-252-991A-32321 | Sequence 32321, A |
| 32 | 55 | 47.4 | 2732 | 4 | US-09-086-436-30 | Sequence 30, Appl |
| 33 | 54.5 | 47.0 | 801 | 1 | US-07-906-349A-6 | Sequence 6, Appl |
| 34 | 54 | 46.6 | 18 | 3 | US-09-604-864-8 | Sequence 8, Appl |
| 35 | 54 | 46.6 | 18 | 3 | US-09-604-864-9 | Sequence 9, Appl |
| 36 | 54 | 46.6 | 18 | 4 | US-10-042-872-8 | Sequence 8, Appl |
| 37 | 54 | 46.6 | 18 | 4 | US-10-042-872-9 | Sequence 9, Appl |
| 38 | 53.5 | 46.1 | 1917 | 4 | US-09-627-650B-5 | Sequence 5, Appl |
| 39 | 53.5 | 46.1 | 1917 | 4 | US-09-436-063C-5 | Sequence 5, Appl |
| 40 | 53 | 45.7 | 1652 | 4 | US-09-627-650B-1 | Sequence 1, Appl |
| 41 | 53 | 45.7 | 1652 | 4 | US-09-436-063C-1 | Sequence 1, Appl |
| 42 | 53 | 45.7 | 2508 | 4 | US-09-627-650B-7 | Sequence 7, Appl |
| 43 | 53 | 45.7 | 2508 | 4 | US-09-436-063C-7 | Sequence 7, Appl |
| 44 | 53 | 45.7 | 2544 | 4 | US-09-627-650B-3 | Sequence 3, Appl |
| 45 | 53 | 45.7 | 2544 | 4 | US-09-436-063C-3 | Sequence 3, Appl |

ALIGNMENTS

RESULT 1
US-09-309-487-1
; Sequence 1, Application US/09309487
; Patent No. 6335318
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/309,487
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-309-487-1

Query Match 87.1%; Score 101; DB 3; Length 18;
Best Local Similarity 83.3%; Pred. No. 1.3e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 GFCRCICTRGFCRCICTR 18
DB 1 GFCRCICTRGFCRCICTR 18

RESULT 2
US-09-967-808-1
; Sequence 1, Application US/09967808
; Patent No. 6514727
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/967,808
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 18

Db 3 GFCRCICRRGVCRCIC 18

RESULT 7

US-10-141-645-1

; Sequence 1, Application US/10141645

; Patent No. 6713078

; GENERAL INFORMATION:

; APPLICANT: Robert Lehrer

; APPLICANT: Alan Waring

; APPLICANT: Alexander Cole

; APPLICANT: Teresa Hong

; TITLE OF INVENTION: Retrocyclins - Antiviral and

; FILE REFERENCE: UCLA-001CIP

; CURRENT APPLICATION NUMBER: US/10/141,645

; CURRENT FILING DATE: 2002-05-06

; PRIOR APPLICATION NUMBER: 60/284,855

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: Unassigned

; NUMBER OF SEQ ID NOS: 125

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-141-645-1

Query Match 77.6%; Score 90; DB 4; Length 18;

Best Local Similarity 77.8%; Pred. No. 0.00022; Indels 4; Gaps 0;

Matches 14; Conservative 0; Mismatches 4; Indels 4; Gaps 0;

QY 1 GFCRCICRRGVCRCIC 18

Db 1 GFCRCICRRGVCRCIC 18

RESULT 8

US-09-917-340-53

; Sequence 53, Application US/09917340

; Patent No. 6696238

; GENERAL INFORMATION:

; APPLICANT: Murphy, Christopher J.

; APPLICANT: McAnulty, Jonathan F.

; APPLICANT: Reid, Ted W.

; TITLE OF INVENTION: Transplant Media

; FILE REFERENCE: TPLANT-06468

; CURRENT APPLICATION NUMBER: US/09/917,340

; CURRENT FILING DATE: 2001-07-29

; PRIOR APPLICATION NUMBER: 60/221,632

; PRIOR FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: 60/249,602

; PRIOR FILING DATE: 2000-11-17

; PRIOR APPLICATION NUMBER: 60/290,932

; PRIOR FILING DATE: 2001-05-15

; NUMBER OF SEQ ID NOS: 96

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 53

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Macaca mulatta

US-09-917-340-53

Query Match 75.0%; Score 87; DB 4; Length 18;

Best Local Similarity 86.7%; Pred. No. 0.00047; Indels 1; Gaps 0;

Matches 13; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 4 RCICRRGVCRCIC 18

Db 1 RCICRRGVCRCIC 15

RESULT 9

US-10-141-645-2

; Sequence 2, Application US/10141645

; Patent No. 6713078

; GENERAL INFORMATION:

; APPLICANT: Robert Lehrer

; APPLICANT: Alan Waring

; APPLICANT: Alexander Cole

; APPLICANT: Teresa Hong

; TITLE OF INVENTION: Retrocyclins - Antiviral and

; FILE REFERENCE: UCLA-001CIP

; CURRENT APPLICATION NUMBER: US/10/141,645

; CURRENT FILING DATE: 2002-05-06

; PRIOR APPLICATION NUMBER: 60/284,855

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: Unassigned

; NUMBER OF SEQ ID NOS: 125

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: synthetic variant

US-10-141-645-2

Query Match 75.0%; Score 87; DB 4; Length 18;

Best Local Similarity 72.2%; Pred. No. 0.00047; Indels 4; Gaps 0;

Matches 13; Conservative 1; Mismatches 4; Indels 4; Gaps 0;

QY 1 GFCRCICRRGVCRCIC 18

Db 1 GFCRCICRRGVCRCIC 18

RESULT 10

US-10-141-645-3

; Sequence 3, Application US/10141645

; Patent No. 6713078

; GENERAL INFORMATION:

; APPLICANT: Robert Lehrer

; APPLICANT: Alan Waring

; APPLICANT: Alexander Cole

; APPLICANT: Teresa Hong

; TITLE OF INVENTION: Retrocyclins - Antiviral and

; FILE REFERENCE: UCLA-001CIP

; CURRENT APPLICATION NUMBER: US/10/141,645

; CURRENT FILING DATE: 2002-05-06

; PRIOR APPLICATION NUMBER: 60/284,855

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: Unassigned

; NUMBER OF SEQ ID NOS: 125

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: synthetic variant

US-10-141-645-3

Query Match 73.3%; Score 85; DB 4; Length 18;

Best Local Similarity 72.2%; Pred. No. 0.00079; Indels 5; Gaps 0;

Matches 13; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 1 GFCRCICRRGVCRCIC 18

Db 1 GFCRCICRRGVCRCIC 18

RESULT 11
US-10-141-645-4
; Sequence 4, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; TITLE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic variant
US-10-141-645-4

Query Match 73.3%; Score 85; DB 4; Length 18;
Best Local Similarity 72.2%; Pred. No. 0.00079;
Matches 13; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | |
Db 1 GICRCICGCGICRCICGR 18

RESULT 12
US-10-141-645-7
; Sequence 7, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; TITLE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic variant
US-10-141-645-7

Query Match 71.6%; Score 83; DB 4; Length 18;
Best Local Similarity 72.2%; Pred. No. 0.0013;
Matches 13; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | |

Db 1 GICYICGGICRCICGR 18

RESULT 13
US-10-141-645-8
; Sequence 8, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; TITLE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic variant
US-10-141-645-8

Query Match 64.7%; Score 75; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.01;
Matches 12; Conservative 0; Mismatches 6; Indels 6; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | |
Db 1 GICYICGGICRCICGR 18

RESULT 14
US-10-141-645-9
; Sequence 9, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; TITLE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic variant
US-10-141-645-9

Query Match 64.7%; Score 75; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.01;
Matches 12; Conservative 0; Mismatches 6; Indels 6; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18

Db 1 GICICGCGICGICG 18

RESULT 15

US-09-309-487-21
 ; Sequence 21, Application US/09309487
 ; Patent No. 6335318
 ; GENERAL INFORMATION:
 ; APPLICANT: Selsted, Michael E.
 ; APPLICANT: Tang, Yi-Quan
 ; APPLICANT: Yuan, Jun
 ; APPLICANT: Ouellette, Andre J.
 ; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
 ; FILE REFERENCE: P-UC 3095
 ; CURRENT APPLICATION NUMBER: US/09/309,487
 ; CURRENT FILING DATE: 1999-05-10
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 21
 ; LENGTH: 92
 ; TYPE: PRT
 ; ORGANISM: Macaca mulatta
 US-09-309-487-21

Query Match 62.9%; Score 73; DB 3; Length 92;
 Best Local Similarity 84.8%; Pred. No. 0.062;
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCICTRGFCRCIC 16
 ||||| :
 Db 65 RCICTRGFCRLC 77

RESULT 16

US-09-967-808-21
 ; Sequence 21, Application US/09967808
 ; Patent No. 6514727
 ; GENERAL INFORMATION:
 ; APPLICANT: Selsted, Michael E.
 ; APPLICANT: Tang, Yi-Quan
 ; APPLICANT: Yuan, Jun
 ; APPLICANT: Ouellette, Andre J.
 ; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
 ; FILE REFERENCE: P-UC 3095
 ; CURRENT APPLICATION NUMBER: US/09/967,808
 ; CURRENT FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: US/09/309,487
 ; PRIOR FILING DATE: 1999-05-10
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 21
 ; LENGTH: 92
 ; TYPE: PRT
 ; ORGANISM: Macaca mulatta
 US-09-967-808-21

Query Match 62.9%; Score 73; DB 4; Length 92;
 Best Local Similarity 84.8%; Pred. No. 0.062;
 Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCICTRGFCRCIC 16
 ||||| :
 Db 65 RCICTRGFCRLC 77

RESULT 17

US-09-309-487-14
 ; Sequence 14, Application US/09309487
 ; Patent No. 6335318
 ; GENERAL INFORMATION:
 ; APPLICANT: Selsted, Michael E.

; APPLICANT: Tang, Yi-Quan
 ; APPLICANT: Yuan, Jun
 ; APPLICANT: Ouellette, Andre J.
 ; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
 ; FILE REFERENCE: P-UC 3095
 ; CURRENT APPLICATION NUMBER: US/09/309,487
 ; CURRENT FILING DATE: 1999-05-10
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 14
 ; LENGTH: 76
 ; TYPE: PRT
 ; ORGANISM: Macaca mulatta
 US-09-309-487-14

Query Match 55.2%; Score 64; DB 3; Length 76;
 Best Local Similarity 83.3%; Pred. No. 0.54;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCICTRGFCRCI 15
 ||||| :
 Db 65 RCICTRGFCRL 76

RESULT 18

US-09-309-487-29
 ; Sequence 29, Application US/09309487
 ; Patent No. 6335318
 ; GENERAL INFORMATION:
 ; APPLICANT: Selsted, Michael E.
 ; APPLICANT: Tang, Yi-Quan
 ; APPLICANT: Yuan, Jun
 ; APPLICANT: Ouellette, Andre J.
 ; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
 ; FILE REFERENCE: P-UC 3095
 ; CURRENT APPLICATION NUMBER: US/09/309,487
 ; CURRENT FILING DATE: 1999-05-10
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 29
 ; LENGTH: 76
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-309-487-29

Query Match 55.2%; Score 64; DB 3; Length 76;
 Best Local Similarity 83.3%; Pred. No. 0.54;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCICTRGFCRCI 15
 ||||| :
 Db 65 RCICTRGFCRL 76

RESULT 19

US-09-967-808-14
 ; Sequence 14, Application US/09967808
 ; Patent No. 6514727
 ; GENERAL INFORMATION:
 ; APPLICANT: Selsted, Michael E.
 ; APPLICANT: Tang, Yi-Quan
 ; APPLICANT: Yuan, Jun
 ; APPLICANT: Ouellette, Andre J.
 ; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
 ; FILE REFERENCE: P-UC 3095
 ; CURRENT APPLICATION NUMBER: US/09/967,808
 ; CURRENT FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: US/09/309,487
 ; PRIOR FILING DATE: 1999-05-10
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 14

us-10-009-317a-32.ra1

Wed Oct 27 08:15:52 2004

```

; LENGTH: 76
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-967-808-14

Query Match      55.2%; Score 64; DB 4; Length 76;
Best Local Similarity 83.3%; Pred. No. 0.54;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 RCICTRGFCRCI 15
Db      65 RCICTRGFCRL 76

RESULT 20
US-09-967-808-29
; Sequence 29, Application US/09967808
; Patent No. 6514727
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/967,808
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 29
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-967-808-29

Query Match      55.2%; Score 64; DB 4; Length 76;
Best Local Similarity 83.3%; Pred. No. 0.54;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 RCICTRGFCRCI 15
Db      65 RCICTRGFCRL 76

RESULT 21
US-10-141-645-15
; Sequence 15, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; TITLE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-00ICIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Macaca mulatta
; FEATURE:
; NAME/KEY: SIGNAL
US-09-967-808-14

Query Match      55.2%; Score 64; DB 4; Length 76;
Best Local Similarity 83.3%; Pred. No. 0.54;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 RCICTRGFCRCI 15
Db      65 RCICTRGFCRL 76

RESULT 22
US-09-510-238A-286
; Sequence 286, Application US/09510238A
; Patent No. 6730303
; GENERAL INFORMATION:
; APPLICANT: FENG, YIQING
; APPLICANT: BAUM, CHARLES M
; APPLICANT: CAPARON, MAIRE H
; APPLICANT: ZURFLUH, LINDA L
; APPLICANT: KLEIN, BARBARA K
; APPLICANT: MCWHERTER, CHARLES A
; APPLICANT: STATEN, NICHOLAS R
; APPLICANT: SUMMERS, NEENA L
; APPLICANT: BAUER, S C
; APPLICANT: LEE, STEPHEN C
; TITLE OF INVENTION: MULTI-FUNCTIONAL HEMATOPOIETIC RECEPTOR
; NUMBER OF SEQUENCES: 313
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DENNIS A. BENNETT, G.D. SEARLE & CO.,
; STREET: P.O. BOX 5110
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/510,238A
; FILING DATE: 22-Feb-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US 96/15774
; FILING DATE: 06-OCT-1996
; APPLICATION NUMBER: US 60/004,834
; FILING DATE: 05-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BENNETT, DENNIS A
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: 2910/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-694-5402
; TELEFAX: 314-694-9095
; INFORMATION FOR SEQ ID NO: 286:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 286:
US-09-510-238A-286

Query Match      51.7%; Score 60; DB 4; Length 180;
Best Local Similarity 56.2%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```
QY      1  GPCRCTGRCGRCIC 16
Db      8  GACCCACTGGCTCTC 23

RESULT 23
US-09-309-487-18
; Sequence 18, Application US/09309487
; Patent No. 6335318
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/309,487
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-309-487-18

Query Match      50.0%; Score 58; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4  RCICTRGFC 12
Db      1  RCICTRGFC 9

RESULT 24
US-09-967-808-18
; Sequence 18, Application US/09967808
; Patent No. 6514727
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
; FILE REFERENCE: Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/967,808
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-967-808-18

Query Match      50.0%; Score 58; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4  RCICTRGFC 12
Db      1  RCICTRGFC 9

RESULT 25
US-10-141-645-36
; Sequence 36, Application US/10141645
; Patent No. 6713078
```

```
QY      1  GPCRCTGRCGRCIC 16
Db      8  GACCCACTGGCTCTC 23

GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; TITLE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: generated by replacement of variants in consensus
; OTHER INFORMATION: sequence
US-10-141-645-36

Query Match      50.0%; Score 58; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4  RCICTRGFC 12
Db      1  RCICTRGFC 9

RESULT 26
US-10-141-645-49
; Sequence 49, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; TITLE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: generated by replacement of variants in consensus
; OTHER INFORMATION: sequence
US-10-141-645-49

Query Match      48.3%; Score 56; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 3.8e+05;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      4  RCICTRGFC 12
Db      1  RCICTRGFC 9

RESULT 27
US-10-141-645-49
```


; Sequence 32502, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32502
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32502

Query Match 47.4%; Score 55; DB 4; Length 168;
Best Local Similarity 61.5%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CICTRGFCRCICT 17
Db 150 CRCSRGFCRCACS 162

RESULT 31

US-09-252-991A-32321
; Sequence 32321, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32321
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32321

Query Match 47.4%; Score 55; DB 4; Length 585;
Best Local Similarity 47.8%; Pred. No. 27;
Matches 11; Conservative 2; Mismatches 4; Indels 6; Gaps 1;

Qy 1 GFCRCICTRG-----FCRCICT 17
Db 532 GRCRALGTRGGKGVGCRCLCT 554

RESULT 32

US-09-086-436-30
; Sequence 30, Application US/09086436
; Patent No. 6703485
; GENERAL INFORMATION:
; APPLICANT: Kandel, Eric R.
; APPLICANT: Santoro, Bina
; APPLICANT: Bartsch, Susan
; APPLICANT: Siegelbaum, Steven
; APPLICANT: Tibbs, Gareth
; APPLICANT: Grant, Seth
; TITLE OF INVENTION: Brain or Heart Cyclic Nucleotide Gated Ion Channel and
; TITLE OF INVENTION: Uses Thereof

; FILE REFERENCE: 0575/54806-A
; CURRENT APPLICATION NUMBER: US/09/086,436
; CURRENT FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 2732
; TYPE: PRT
; ORGANISM: Murine
US-09-086-436-30

Query Match 47.4%; Score 55; DB 4; Length 2732;
Best Local Similarity 57.1%; Pred. No. 90;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 CRICIRGFCRCIC 16
Db 2096 CCCCCTGCCACAC 2109

RESULT 33

US-07-906-349A-6
; Sequence 6, Application US/07906349A
; Patent No. 5434064
; GENERAL INFORMATION:
; APPLICANT: Schllessinger, Joseph
; APPLICANT: Skolnik, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AND
; TITLE OF INVENTION: TARGET PROTEINS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,349A
; FILING DATE: 30-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/643,237
; FILING DATE: 18-JAN-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 801 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-07-906-349A-6

Query Match 47.0%; Score 54.5; DB 1; Length 801;
Best Local Similarity 55.6%; Pred. No. 39;
Matches 10; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 1 GFCRCICT-RGFCRCICT 17
Db 345 GTCACCCCTGTGACCCCTCT 362

RESULT 34

US-09-604-864-8
 ; Sequence 8, Application US/09604864
 ; Patent No. 6337317
 ; GENERAL INFORMATION:
 ; APPLICANT: Hancock, Robert E. W.
 ; APPLICANT: Zhang, Lijuan
 ; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES AND METHODS OF
 ; FILE REFERENCE: UBC1170
 ; CURRENT APPLICATION NUMBER: US/09/604,864
 ; CURRENT FILING DATE: 2000-06-27
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Limulus polyphemus
 US-09-604-864-8

Query Match 46.6%; Score 54; DB 3; Length 18;
 Best Local Similarity 60.0%; Pred. No. 2.4;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 FRCICITRGFCRCIC 16
 :||| ||||| |
 Db 3 WCRRCYRGFCRRFC 17

RESULT 35

US-09-604-864-9
 ; Sequence 9, Application US/09604864
 ; Patent No. 6337317
 ; GENERAL INFORMATION:
 ; APPLICANT: Hancock, Robert E. W.
 ; APPLICANT: Zhang, Lijuan
 ; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES AND METHODS OF
 ; FILE REFERENCE: UBC1170
 ; CURRENT APPLICATION NUMBER: US/09/604,864
 ; CURRENT FILING DATE: 2000-06-27
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Limulus polyphemus
 US-09-604-864-9

Query Match 46.6%; Score 54; DB 3; Length 18;
 Best Local Similarity 60.0%; Pred. No. 2.4;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 FRCICITRGFCRCIC 16
 :||| ||||| |
 Db 3 WCRRCYRGFCRRFC 17

RESULT 36

US-10-042-872-8
 ; Sequence 8, Application US/10042872
 ; Patent No. 6747007
 ; GENERAL INFORMATION:
 ; APPLICANT: THE UNIVERSITY OF BRITISH COLUMBIA
 ; APPLICANT: Hancock, Robert E. W.
 ; APPLICANT: Zhang, Lijuan
 ; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES AND METHODS OF
 ; FILE REFERENCE: UBC1170-1
 ; CURRENT APPLICATION NUMBER: US/10/042,872
 ; CURRENT FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 09/604,864
 ; PRIOR FILING DATE: 2000-06-27
 ; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Limulus polyphemus
 US-10-042-872-8

Query Match 46.6%; Score 54; DB 4; Length 18;
 Best Local Similarity 60.0%; Pred. No. 2.4;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 FRCICITRGFCRCIC 16
 :||| ||||| |
 Db 3 WCRRCYRGFCRRFC 17

RESULT 37

US-10-042-872-9
 ; Sequence 9, Application US/10042872
 ; Patent No. 6747007
 ; GENERAL INFORMATION:
 ; APPLICANT: THE UNIVERSITY OF BRITISH COLUMBIA
 ; APPLICANT: Hancock, Robert E. W.
 ; APPLICANT: Zhang, Lijuan
 ; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES AND METHODS OF
 ; FILE REFERENCE: UBC1170-1
 ; CURRENT APPLICATION NUMBER: US/10/042,872
 ; CURRENT FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 09/604,864
 ; PRIOR FILING DATE: 2000-06-27
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Limulus polyphemus
 US-10-042-872-9

Query Match 46.6%; Score 54; DB 4; Length 18;
 Best Local Similarity 60.0%; Pred. No. 2.4;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 FRCICITRGFCRCIC 16
 :||| ||||| |
 Db 3 WCRRCYRGFCRRFC 17

RESULT 38

US-09-627-650B-5
 ; Sequence 5, Application US/09627650B
 ; Patent No. 6406872
 ; GENERAL INFORMATION:
 ; APPLICANT: Bamber, Bruce
 ; APPLICANT: Jorgensen, Erik
 ; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
 ; FILE REFERENCE: 21101.000903
 ; CURRENT APPLICATION NUMBER: US/09/627,650B
 ; CURRENT FILING DATE: 2000-07-28
 ; PRIOR APPLICATION NUMBER: 09/436,063
 ; PRIOR FILING DATE: 1999-11-08
 ; PRIOR APPLICATION NUMBER: 60/107,727
 ; PRIOR FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 1917
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-09-627-650B-5

Query Match 46.1%; Score 53.5; DB 4; Length 1917;

Best Local Similarity 60.0%; Pred. No. 1e+02; Mismatches 0; Indels 5; Gaps 1;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 3 CRCTCTGFCRCICT 17
Db 67 CTCTCT-GTCTCTCT 80

RESULT 39

US-09-436-063C-5
; Sequence 5, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematoe Neuromuscular Junction GABA Receptors and
; TITLE OF INVENTION: Methods Related Thereto
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1917
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-436-063C-5

Query Match 46.1%; Score 53.5; DB 4; Length 1917;
Best Local Similarity 60.0%; Pred. No. 1e+02; Mismatches 0; Indels 5; Gaps 1;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 3 CRCTCTGFCRCICT 17
Db 67 CTCTCT-GTCTCTCT 80

RESULT 40

US-09-627-650B-1
; Sequence 1, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematoe Neuromuscular Junction GABA Receptors and
; TITLE OF INVENTION: Methods Related Thereto
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1652
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-627-650B-1

Query Match 45.7%; Score 53; DB 4; Length 1652;
Best Local Similarity 47.1%; Pred. No. 1e+02; Mismatches 0; Indels 9; Gaps 0;
Matches 8; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICT 17
Db 232 GACTCAGCATCGCACT 248

Search completed: October 26, 2004, 15:41:18

OM protein - protein search, using sw model

Run on: October 26, 2004, 15:15:31 ; Search time 23.5 Seconds
(without alignments)
73.698 Million cell updates/sec

Title: US-10-009-317A-32

Perfect score: 116

Sequence: 1 GFCRCICTRGFCRCICTR 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 87 | 75.0 | 18 | 2 C59089 | theta defensin-1 - |
| 2 | 64 | 55.2 | 76 | 2 A59089 | theta defensin la |
| 3 | 54.5 | 47.0 | 248 | 2 E71602 | probable integral |
| 4 | 54 | 46.6 | 290 | 2 G72858 | AcOrf-70 protein - |
| 5 | 53 | 45.7 | 164 | 2 T24272 | hypothetical prote |
| 6 | 53 | 45.7 | 188 | 2 T15651 | hypothetical prote |
| 7 | 52.5 | 45.3 | 73 | 2 JC1066 | trypsin inhibitor |
| 8 | 51 | 44.0 | 72 | 1 TIMB | trypsin inhibitor |
| 9 | 51 | 44.0 | 419 | 2 S69207 | vascular endotheli |
| 10 | 50.5 | 43.5 | 83 | 2 S07405 | proteinase inhibit |
| 11 | 50.5 | 43.5 | 94 | 2 JC2225 | Bowman-Birk protei |
| 12 | 50.5 | 43.5 | 103 | 1 T1SYC2 | proteinase inhibit |
| 13 | 50 | 43.1 | 624 | 1 S54581 | probable membrane |
| 14 | 49 | 42.2 | 152 | 2 T18975 | hypothetical prote |
| 15 | 49 | 42.2 | 157 | 2 A25964 | thyroglobulin - ra |
| 16 | 49 | 42.2 | 170 | 2 S08167 | Balbain ring 3 pr |
| 17 | 48 | 41.4 | 72 | 1 T12B1P | proteinase inhibit |
| 18 | 48 | 41.4 | 78 | 1 T12B1A | proteinase inhibit |
| 19 | 48 | 41.4 | 79 | 1 T1F82 | proteinase inhibit |
| 20 | 48 | 41.4 | 83 | 2 T07941 | proteinase inhibit |
| 21 | 48 | 41.4 | 102 | 1 T1SYD2 | proteinase inhibit |
| 22 | 48 | 41.4 | 160 | 2 T25185 | proteinase inhibit |
| 23 | 48 | 41.4 | 306 | 2 S32834 | hypothetical prote |
| 24 | 48 | 41.4 | 476 | 2 JC5042 | G protein-coupled |
| 25 | 48 | 41.4 | 1353 | 1 JH0675 | methyloxygen-red |
| 26 | 47.5 | 40.9 | 77 | 2 I48725 | restrictin precurs |
| 27 | 47.5 | 40.9 | 550 | 2 PQ0618 | Q300 protein - mou |
| 28 | 47 | 40.5 | 113 | 2 S56648 | hypothetical prote |
| 29 | 47 | 40.5 | 317 | 2 JC7597 | trypsin inhibitor |
| | | | | | chondromodulin-I 1 |

| | | | | | |
|----|------|------|------|----------|--------------------|
| 30 | 47 | 40.5 | 317 | 2 JC7603 | tenomodulin - mous |
| 31 | 47 | 40.5 | 2195 | 2 T34264 | hypothetical prote |
| 32 | 47 | 40.5 | 2946 | 2 T15840 | hypothetical prote |
| 33 | 47 | 40.5 | 3191 | 2 T22945 | hypothetical prote |
| 34 | 47 | 40.5 | 4660 | 2 T42737 | gp330 protein prec |
| 35 | 46.5 | 40.1 | 141 | 2 T33983 | hypothetical prote |
| 36 | 46.5 | 40.1 | 883 | 2 T49781 | related to mutanas |
| 37 | 46.5 | 40.1 | 3461 | 2 S58870 | zeelin precursor - |
| 38 | 46 | 39.7 | 146 | 2 S05415 | proteinase inhibit |
| 39 | 46 | 39.7 | 511 | 2 T17298 | hypothetical prote |
| 40 | 46 | 39.7 | 761 | 2 T03052 | hypothetical prote |
| 41 | 46 | 39.7 | 792 | 2 C96719 | hypothetical prote |
| 42 | 46 | 39.7 | 1251 | 2 A57293 | latent proteinin |
| 43 | 46 | 39.7 | 1307 | 2 G96711 | unknown protein, 9 |
| 44 | 46 | 39.7 | 1408 | 2 S16148 | gene serrate prote |
| 45 | 45.5 | 39.2 | 222 | 2 A31685 | EFl protein - fowl |

ALIGNMENTS

RESULT 1

C59089
theta defensin-1 - rhesus macaque
N/Alternate names: RTD-1
C/Species: Macaca mulatta (rhesus macaque)
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C/Accession: C59089
R/Tang, Y.Q.; Yuan, J.; Osapay, G.; Osapay, K.; Tran, D.; Miller, C.J.; Ouellette, A.J.;
Science 286, 498-502, 1999
A/Title: A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of
F1-9/Region: theta defensin la-derived
A/Reference number: A59089; MUID:99453140; PMID:10521339
A/Accession: C59089
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-18 <SEL>
A/Note: this sequence is cyclically permuted by -6 residues from the sequence presented i
C/Comment: For the two contributing precursor sequences, see PIR:A59089 and PIR:B59089.
C/Keywords: antibacterial; antibiotic; antifungal; leukocyte; protein splicing
F1-9/Region: theta defensin la-derived
F10-18/Region: theta defensin lb-derived
F1-18/Cross-link: cyclopeptide (Arg-Cys) #status experimental
F2-11,4-9,13-18/Disulfide bonds: #status experimental
F19-10/Cross-link: cyclopeptide (Cys-Arg) #status experimental

Query Match 75.0%; Score 87.0; DB 2; Length 18;
Best Local Similarity 86.7%; Pred. No. 0.00014;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCICTRGFCRCICTR 18
||| ||||| |||
Db 1 RCICTRGFCRCICTR 15

RESULT 2

A59089
theta defensin la precursor - rhesus macaque
C/Species: Macaca mulatta (rhesus macaque)
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: A59089
R/Tang, Y.Q.; Yuan, J.; Osapay, G.; Osapay, K.; Tran, D.; Miller, C.J.; Ouellette, A.J.;
Science 286, 498-502, 1999
A/Title: A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of
F1-9/Region: theta defensin la-derived
A/Reference number: A59089; MUID:99453140; PMID:10521339
A/Accession: A59089
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-76 <TAN>
A/Cross-references: UNIPROT:P82270; GB:AF191100; NID:96137227; PIDN:AAF04389.1; PID:96137
C/Comment: For the complete mature sequence, see PIR:C59089.
C/Keywords: antibacterial; antibiotic; antifungal; leukocyte; protein splicing
F1-20/Domain: signal sequence #status predicted <SIG>

F:21-64/Domain: amino-terminal propeptide #status predicted <PRO>
F:74-76/Domain: carboxyl-terminal propeptide #status predicted <CTP>

Query Match 55.2%; Score 64; DB 2; Length 76;
Best Local Similarity 83.3%; Pred. No. 0.2;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCICTRGFRCICI 15
| | | | | :
Db 65 RCICTRGFRCICI 76

RESULT 3

E71602
probable integral membrane protein PFB0950w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C:Accession: E71602
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Perle, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: E71602
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-248 <GAR>
A:Cross-references: UNIPROT:O96282; GB:AE001428; GB:AE001362; NID:G3845316; PIDN:AA07197
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0950w

Query Match 47.0%; Score 54.5; DB 2; Length 248;
Best Local Similarity 60.0%; Pred. No. 6;
Matches 9; Conservative 1; Mismatches 2; Indels 3; Gaps 1;

QY 3 CRCICTRGFRCICIT 17
| | | | | :
Db 163 CSCICIT---CTGICS 174

RESULT 4

G72858
AcOrf-70 protein - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus
A:Note: dsDNA virus
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 09-Jul-2004
C:Accession: G72858
R:Avres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 588-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.
A:Reference number: A72850; MUID:94303173; PMID:8030224
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-290 <AYR>
A:Cross-references: UNIPROT:P41470; GB:L22858; NID:G510708; PIDN:AAA66700.1; PID:G559139
C:Genetics:
A:Gene: AcOrf-70

Query Match 46.6%; Score 54; DB 2; Length 290;
Best Local Similarity 40.7%; Pred. No. 7;
Matches 11; Conservative 2; Mismatches 4; Indels 10; Gaps 2;

QY 2 FCRICIT-----RGF-CRCICITR 18
| | | | | :
Db 218 FARCFCTNTWQCFRCRQYKCEICIR 244

RESULT 5

T24272
hypothetical protein T01B7.8 - Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24272

R:Sim, M.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z19867
A:Accession: T24272
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-164 <WIL>
A:Cross-references: UNIPROT:Q22048; EMBL:Z66499; PIDN:CAA91301.1; GSPDB:GN00020; CESP:T01
A:Experimental source: clone T01B7
C:Genetics:
A:Gene: CESP:T01B7.8
A:Map position: 2
A:Introns: 20/3; 90/2

Query Match 45.7%; Score 53; DB 2; Length 164;
Best Local Similarity 50.0%; Pred. No. 6.9;
Matches 9; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GPCRICITRGFCRCICITR 18
| | | | | :
Db 80 GCGGCCCCRCRCRCRCRR 97

RESULT 6

T15651
hypothetical protein C27A2.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T15651
R:Nhan, M.

submitted to the EMBL Data Library, May 1996
A:Description: The sequence of C. elegans cosmid C27A2.
A:Reference number: Z18382
A:Accession: T15651
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-188 <NHA>
A:Cross-references: UNIPROT:Q18238; EMBL:U58760; NID:G1330384; PID:G1330389; PIDN:AA00071
A:Experimental source: strain Bristol N2; clone C27A2
C:Genetics:
A:Gene: CESP:C27A2.5
A:Map position: 2
A:Introns: 19/3; 91/2

Query Match 45.7%; Score 53; DB 2; Length 188;
Best Local Similarity 50.0%; Pred. No. 7.6;
Matches 9; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GPCRICITRGFCRCICITR 18
| | | | | :
Db 81 GCGGCCCCRCRCRCRCRR 98

RESULT 7

JC1066
trypsin inhibitor - mung bean
C:Species: Vigna radiata (mung bean)
C:Date: 02-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 18-Aug-2000
C:Accession: JC1066
R:Chen, C.Q.; Mao, J.F.; Zhang, M.F.; Dai, J.F.
Chinese J. Biotechnol. 9, 54-60, 1993
A:Title: Synthesis of mung bean trypsin inhibitor by the combination of the single strand

A:Reference number: JC1066
A:Accession: JC1066
A:Molecule type: DNA
A:Residues: 1-73 <CHE>
A:Superfamily: Bowman-Birk proteinase inhibitor; Bowman-Birk inhibitor repeat homology
C:Keywords: serine proteinase inhibitor
F:2-73/Product: trypsin inhibitor #status predicted <NAT>
F:14-40/Domain: Bowman-Birk inhibitor repeat homology <BB1>
F:41-66/Domain: Bowman-Birk inhibitor repeat homology <BB12>

Query Match 45.3%; Score 52.5; DB 2; Length 73;
Best Local Similarity 48.0%; Pred. No. 4.6;
Matches 12; Conservative 3; Mismatches 1; Indels 9; Gaps 3;

QY 3 CR-CICTR---GFCRCI-----CTR 18
|:|||||:|||||:|
Db 41 CKSCICTRSMFGKRCRLDTRDDFCYK 65
|:|||||:|||||:|

RESULT 8
TIME
trypsin inhibitor (Bowman-Birk) - mung bean
C;Species: Vigna radiata (mung bean)
C;Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 09-Jul-2004
C;Accession: A01301
R;Zhang, Y.; Luo, S.; Tan, F.; Qi, Z.; Xu, L.; Zhang, A.
Sci. Sin. 25, 268-277, 1982
A;Title: Complete amino acid sequence of mung bean trypsin inhibitor.
A;Reference number: A01301
A;Accession: A01301
A;Molecule type: protein
A;Residues: 1-72 <ZHA>
A;Cross-references: UNIPROT:P01062
A;Note: three isoforms are also found whose amino ends differ slightly from that sh
d Lys-1, respectively
C;Comment: This inhibitor stoichiometrically inhibits trypsin in a molar ratio of 1:2.
C;Superfamily: Bowman-Birk proteinase inhibitor; Bowman-Birk inhibitor repeat homology
C;Keywords: duplication; serine proteinase inhibitor
F;13-39/Domain: Bowman-Birk inhibitor repeat homology <BB1>
F;40-65/Domain: Bowman-Birk inhibitor repeat homology <BB2>
F;2-66,13-28,16-62,18-26,36-43,40-55,45-53/Disulfide bonds: #status predicted
F;50/Inhibitory site: Lys (trypsin) #status predicted
F;47/Inhibitory site: Arg (trypsin) #status predicted

Query Match 44.0%; Score 51; DB 1; Length 72;
Best Local Similarity 57.9%; Pred. No. 6.8;
Matches 11; Conservative 2; Mismatches 2; Indels 4; Gaps 2;

QY 3 CR-CICTR---GFCRCICT 17
|:|||||:|||||:|
Db 40 CKSCICTRSMFGKRCRLDTR 58
|:|||||:|||||:|

RESULT 9
TIME
vascular endothelial growth factor C precursor - human
N;Alternate names: FLT4 ligand DHM
C;Species: Homo sapiens (man)
C;Date: 27-Apr-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C;Accession: S69207; S61795; S71443; S69208; G02659
R;Joukov, V.; Pajusola, K.; Kaipainen, A.; Chilov, D.; Lahinen, I.; Kukkk, E.; Saksela,
EMBO J. 15, 1751, 1996
A;Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a ligand fo
A;Reference number: S69207; MUID:96203094; PMID:8612600
A;Accession: S69207
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-419 <JOU>
A;Cross-references: UNIPROT:P49767; EMBL:X94216; NID:g1177488; PIDN:CAA3907.1; PID:e221
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
A;Note: only a part of the translation is shown
R;Joukov, V.; Pajusola, K.; Kaipainen, A.; Chilov, D.; Lahinen, I.; Kukkk, E.; Saksela,
EMBO J. 15, 290-298, 1996
A;Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the Flt4 (V
A;Reference number: S61795; MUID:96178224; PMID:8617204
A;Accession: S61795
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 70-419 <JOU>
A;Note: this sequence has been revised in reference S69207
A;Accession: S71443

A;Gene: CESP:C06A1.6
A;Map position: 2
A;Introns: 22/3

Query Match 42.2%; Score 49; DB 2; Length 152;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GFCRCICRGGFCRCIC 16
Db 67 GCGGCCCCPRCCCC 82

RESULT 15
A25964
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-1989 #sequence_revision 30-Jun-1991 #text_change 11-Apr-1997
C;Accession: A25964
R;Musci, A.M.; Avvedimento, E.V.; Polistina, C.; Urini, V.M.; Obici, S.; Nitech, L.; C
Proc. Natl. Acad. Sci. U.S.A. 83, 323-327, 1986
A;Title: The complete structure of the rat thyroglobulin gene.
A;Reference number: A25964; MUID:86094383; PMID:3455768
A;Accession: A25964
A;Molecule type: DNA
A;Residues: 1-157 <MUS>
A;Cross-references: GB:M12558
A;Note: the authors translated the codon GTG for residue 44 as Leu
C;Superfamily: thyroglobulin; cholinesterase homology; thyroglobulin type I repeat homol
C;Keywords: dimer; glycoprotein; iodine; thyroid gland; thyroid hormone biosynthesis
F;35-60/Domain: thyroglobulin type I repeat homology (fragment) <THY1>
F;134-138/Domain: thyroglobulin type I repeat homology (fragment) <THY2>
F;25/Modified site: thyroxine (Tyr) #status predicted

Query Match 42.2%; Score 49; DB 2; Length 157;
Best Local Similarity 32.0%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 5; Indels 8; Gaps 1;

QY 1 GFCRCIC-----TRGFCRCICT 17
Db 110 GECWCVCSPFVTCTSGECYCVC 134

RESULT 16
S08167
Baibiani ring 3 protein - midge (Chironomus tentans)
C;Species: Chironomus tentans
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Aug-2004
C;Accession: S08167
R;Paulsson, G.; Lendahl, U.; Galli, J.; Ericsson, C.; Wieslander, L.
J. Mol. Biol. 211, 331-349, 1990
A;Title: The Baibiani ring 3 gene in Chironomus tentans has a diverged repetitive struct
A;Reference number: S08167; MUID:90172404; PMID:1689777
A;Accession: S08167
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-1700 <PAU>
A;Cross-references: UNIPROT:Q03376; GB:X52263; NID:g7057; PIDN:CAA36506.1; PID:g7058
C;Genetics:
A;Gene: BR3
A;Map position: 4

Query Match 42.2%; Score 49; DB 2; Length 1700;
Best Local Similarity 23.3%; Pred. No. 1e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 16; Gaps 1;

QY 3 CRCICRGG-----FRCIC 16
Db 1245 CRCVCPRNMEKPADNCKTKWNDMEQCVC 1274

RESULT 17
TIZB1P

protease inhibitor (Bowman-Birk) I-A' - adzuki bean
C;Species: Vigna angularis (adzuki bean)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: C01300; A01300
R;Kiyohara, T.; Yokota, K.; Masaki, Y.; Matsui, O.; Iwasaki, T.; Yoshikawa, M.
J. Biochem. 90, 721-728, 1981
A;Title: The amino acid sequences of proteinase inhibitors I-A and I-A' from adzuki beans
A;Reference number: A01300; MUID:82075699; PMID:7309695
A;Accession: C01300
A;Molecule type: protein
A;Residues: 1-72 <KIY>
A;Cross-references: UNIPROT:P01061
C;Comment: This inhibitor strongly inhibits trypsin.
C;Superfamily: Bowman-Birk proteinase inhibitor; Bowman-Birk inhibitor repeat homology
C;Keywords: duplication; serine proteinase inhibitor
F;13-39/Domain: Bowman-Birk inhibitor repeat homology <BB1>
F;40-65/Domain: Bowman-Birk inhibitor repeat homology <BB2>
F;12-66,13-28,16-62,18-26,36-43,40-55,45-53/Disulfide bonds: #status predicted
F;20/Inhibitory site: Lys (trypsin) #status predicted
F;47/Inhibitory site: Arg (trypsin) #status predicted

Query Match 41.4%; Score 48; DB 1; Length 72;
Best Local Similarity 52.6%; Pred. No. 15;
Matches 10; Conservative 3; Mismatches 2; Indels 4; Gaps 2;

QY 3 CR-CICR---GFCRCICT 17
Db 40 CKSCMCTRSMFGQRCCLDT 58

RESULT 18
TIZB1A
protease inhibitor (Bowman-Birk) I-A - adzuki bean
C;Species: Vigna angularis (adzuki bean)
C;Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 09-Jul-2004
C;Accession: A01300
R;Kiyohara, T.; Yokota, K.; Masaki, Y.; Matsui, O.; Iwasaki, T.; Yoshikawa, M.
J. Biochem. 90, 721-728, 1981
A;Title: The amino acid sequences of proteinase inhibitors I-A and I-A' from adzuki beans
A;Reference number: A01300; MUID:82075699; PMID:7309695
A;Accession: A01300
A;Molecule type: protein
A;Residues: 1-78 <KIY>
A;Cross-references: UNIPROT:P01061
C;Comment: This inhibitor strongly inhibits trypsin.
C;Superfamily: Bowman-Birk proteinase inhibitor; Bowman-Birk inhibitor repeat homology
C;Keywords: duplication; serine proteinase inhibitor
F;19-45/Domain: Bowman-Birk inhibitor repeat homology <BB1>
F;46-71/Domain: Bowman-Birk inhibitor repeat homology <BB2>
F;18-72,19-34,22-68,24-32,42-49,46-61,51-59/Disulfide bonds: #status predicted
F;26/Inhibitory site: Lys (trypsin) #status experimental
F;53/Inhibitory site: Arg (trypsin) #status experimental

Query Match 41.4%; Score 48; DB 1; Length 78;
Best Local Similarity 52.6%; Pred. No. 16;
Matches 10; Conservative 3; Mismatches 2; Indels 4; Gaps 2;

QY 3 CR-CICR---GFCRCICT 17
Db 46 CKSCMCTRSMFGQRCCLDT 64

RESULT 19
TIFB2
proteinase inhibitor (Bowman-Birk) II - kidney bean (tentative sequence)
C;Species: Phaseolus vulgaris (kidney bean)
C;Date: 30-Nov-1980 #sequence_revision 30-Nov-1980 #text_change 09-Jul-2004
C;Accession: A01299
R;Wilson, K.A.; Laskowski Sr., M.
J. Biol. Chem. 250, 4261-4267, 1975
A;Title: The partial amino acid sequence of trypsin inhibitor II from garden bean, Phaseo
A;Reference number: A01299; MUID:75151596; PMID:1126951
A;Accession: A01299

R:Klein, A.
submitted to the EMBL Data Library, August 1991
A:Reference number: S32833
A:Accession: S32834
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-306 <KLE>
A:Cross-references: UNIPROT:Q00406; EMBL:X61203; NID:G296614; PID:G296614
R:Halboth, S.; Klein, A.
submitted to the EMBL Data Library, August 1991
A:Description: Methanococcus voltae harbors two gene groups each of homologous (Nife) - a
A:Reference number: S16721
A:Accession: S16726
A:Molecule type: DNA
A:Residues: 1-99, 'TA', 102-107, 'P', 126, 'LH', 129-306 <HAL>
A:Cross-references: EMBL:X61203
A:Note: The sequence is revised in GenBank entry X61203, release 117, (PIDN:CAA43505.1)
R:Halboth, S.; Klein, A.
Mol. Gen. Genet. 233, 217-224, 1992
A:Title: Methanococcus voltae harbors four gene clusters potentially encoding two [Nife]
A:Reference number: A59304; MUID:9223118; PMID:1603063
A:Contents: annotation
C:Genetics:
A:Gene: vhcG
C:Superfamily: methyl viologen-reducing hydrogenase gamma chain

Query Match 41.4%; Score 48; DB 2; Length 306;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 CRCICTRGFCRCIC 16
Db 215 CMGVATRASRCRCFC 228

RESULT 24

G protein-coupled receptor - barnacle
C:Species: Balanus amphitrite (barnacle)
C:Date: 21-Jan-1997 #sequence_revision 21-Jan-1997 #text_change 09-Jul-2004
A:Accession: JC5042
R:Isai, A.; Kawahara, H.; Okazaki, Y.; Shizuri, Y.
Gene 175, 95-100, 1996
A:Title: Molecular cloning of a new member of the putative G protein-coupled receptor gene
A:Reference number: JC5042; MUID:97074655; PMID:8917082
A:Accession: JC5042
A:Molecule type: DNA
A:Residues: 1-476 <ISO>
A:Cross-references: UNIPROT:Q93126; DDBJ:D78363; NID:g1514430; PIDN:BAAL1375.1; PID:g15144725
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
F:34-57/Domain: transmembrane #status predicted <TM1>
F:71-93/Domain: transmembrane #status predicted <TM2>
F:107-129/Domain: transmembrane #status predicted <TM3>
F:126-134/Region: G protein-binding #status predicted
F:148-172/Domain: transmembrane #status predicted <TM4>
F:199-222/Domain: transmembrane #status predicted <TM5>
F:377-398/Domain: transmembrane #status predicted <TM6>
F:407-430/Domain: transmembrane #status predicted <TM7>
F:13.17/Binding site: carbohydrate (asn) (covalent) #status predicted
F:230,240,250,260/Binding site: phosphate (thr) (covalent) (by protein kinase C) #status predicted
F:355/Binding site: phosphate (ser) (covalent) (by cAMP-dependent kinase) #status predicted

Query Match 41.4%; Score 48; DB 2; Length 476;
Best Local Similarity 26.5%; Pred. No. 55;
Matches 9; Conservative 5; Mismatches 4; Indels 16; Gaps 2;

Qy 1 GPCR-----CICTRGF-----CRCICTR 18
Db 416 GYCNNAVPCVYALCSDFRFAFSSILCKVCRR 449

RESULT 25

JH0675
restrictin precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
A:Accession: JH0675; PS0385; S23254
R:Noerenberg, U.; Wille, H.; Wolff, J.M.; Frank, R.; Rathjen, F.G.
Neuron 8, 849-863, 1992
A:Title: The chicken neural extracellular matrix molecule restrictin: similarity with EGI
A:Reference number: JH0675; MUID:92265298; PMID:1375037
A:Accession: JH0675
A:Molecule type: mRNA
A:Residues: 1-1353 <NOE>
A:Cross-references: UNIPROT:Q00546; GB:X64649; NID:G63613; PIDN:CAA45920.1; PID:G63614
A:Experimental source: brain
A:Accession: PS0385
A:Molecule type: protein
A:Residues: 579-586;827-840 <NOE1>
C:Comment: This protein is a neural extracellular matrix protein implicated in neural cell
C:Superfamily: restrictin; EGF homology; fibronectin beta/gamma homology; fibronectin type
C:Keywords: calcium binding; cell adhesion; duplication; extracellular matrix; glycoprotein
F:1-33/Domain: signal sequence #status predicted <SIG>
F:34-1353/Product: restrictin #status predicted <NAT>
F:203-229/Domain: EGF homology <EG1>
F:234-260/Domain: EGF homology <EG2>
F:265-291/Domain: EGF homology <EG3>
F:296-322/Domain: EGF homology <EG4>
F:324-405/Domain: fibronectin type III repeat homology <FN1>
F:413-494/Domain: fibronectin type III repeat homology <FN2>
F:502-584/Domain: fibronectin type III repeat homology <FN3>
F:592-676/Domain: fibronectin type III repeat homology <FN4>
F:684-764/Domain: fibronectin type III repeat homology <FN5>
F:772-853/Domain: fibronectin type III repeat homology <FN6>
F:861-941/Domain: fibronectin type III repeat homology <FN7>
F:949-1027/Domain: fibronectin type III repeat homology <FN8>
F:1035-1115/Domain: fibronectin type III repeat homology <FN9>
F:1130-1338/Domain: fibronectin beta/gamma homology <FBG>
F:1172-1286/Region: calcium binding #status predicted
F:53,197,277,391,469,580,734,790,960,1031,1041,1256,1342/Binding site: carbohydrate (asn)

Query Match 41.4%; Score 48; DB 1; Length 1353;
Best Local Similarity 46.7%; Pred. No. 1.1e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CRCICTRGFCRCIC 17
Db 185 CRCICSEGWAGNCS 199

RESULT 26

I48725
Q300 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
A:Accession: I48725
R:Wagner, S.; Cullmann, G.; Knippers, R.
J. Virol. 65, 3259-3267, 1991
A:Title: The Q300 gene: a novel transcription unit induced in simian virus 40-infected ar
A:Reference number: I48725; MUID:91237845; PMID:1851876
A:Accession: I48725
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-77 <RES>
A:Cross-references: UNIPROT:Q02722; EMBL:X52164; NID:G53860; PIDN:CAA36417.1; PID:G53861

Query Match 40.9%; Score 47.5; DB 2; Length 77;
Best Local Similarity 46.7%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

Qy 3 CRCICTRGFCRCIC 17
Db 30 CVCVCV---CVCVCT 41

Db 1215 VCVNGFCRC 1223

RESULT 32

T15840

hypothetical protein C54G7.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15840

R:Du, Z.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C54G7.

A:Reference number: Z18416

A:Accession: T15840

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-2946 <DUZ>

A:Cross-references: EMBL:U40410; NID:g1065453; PID:g1065455; PIDN:AAA81392.1; CESP:C54G7.3

C:Gene: CESP:C54G7.3

A:Introns: 16/1; 53/2; 92/1; 150/3; 295/1; 346/1; 392/1; 440/1; 475/3; 579/1; 615/3; 169/46/1; 2561/2; 2603/2; 2626/2; 2665/2; 2716/3; 2804/3; 2884/3

Query Match 40.5%; Score 47; DB 2; Length 2946;

Best Local Similarity 66.7%; Pred. No. 2.5e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ICITRGFCRC 14

Db 2430 VCTNGFCRC 2438

RESULT 33

T22945

hypothetical protein T01D3.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T22945

R:Percy, C.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19640

A:Accession: T22945

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-3191 <WIL>

A:Cross-references: UNIPROT:O01335; EMBL:Z81094; PIDN:CAB03155.1; GSPDB:GN00023; CESP:T01D3.1

A:Experimental source: clone F58G11

R:Steward, C.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19870

A:Accession: T24295

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-3191 <W12>

A:Cross-references: EMBL:Z81110; PIDN:CAB03261.1; GSPDB:GN00023; CESP:T01D3.1

A:Experimental source: clone T01D3

C:Gene: CESP:T01D3.1

A:Map position: 5

A:Introns: 18/1; 73/2; 101/3; 241/2; 251/1; 327/3; 383/3; 455/3; 494/1; 793/1; 905/3; 10

Query Match 40.5%; Score 47; DB 2; Length 3191;

Best Local Similarity 30.6%; Pred. No. 2.6e+02;

Matches 11; Conservative 1; Mismatches 2; Indels 22; Gaps 2;

Qy 1 GFCC-----ICITRGFCRC 14

Db 1182 GLCKEDGQGRSQIPLCNSCSLNGICTRPGFCSC 1217

RESULT 34

T42737

gp330 protein precursor - rat

N:Alternate names: megalin

C:Species: Rattus norvegicus (Norway rat)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004

C:Accession: T42737

R:Saio, A.; Pietromaco, S.; Loo, A.K.C.; Farquhar, M.G.

Proc. Natl. Acad. Sci. U.S.A. 91, 9725-9729, 1994

A:Title: Complete cloning and sequencing of rat gp330/megalin, a distinctive member of the

A:Reference number: A58173; NUID:95024033; PMID:7937880

A:Accession: T42737

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-4660 <SAI>

A:Cross-references: UNIPROT:P98158; EMBL:L34049; NID:g561852; PID:g561853; PIDN:AAA51369.

A:Experimental source: strain Sprague-Dawley; kidney

A:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-4660/Product: gp330 protein #status predicted <MAT>

Query Match 40.5%; Score 47; DB 2; Length 4660;

Best Local Similarity 50.0%; Pred. No. 3.4e+02;

Matches 9; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

Qy 1 GFC--RCITRGFCRC 16

Db 1399 GFCGHCVCNVRGSRFCAC 1416

RESULT 35

T33983

hypothetical protein H23N18.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T33983

R:Courtnay, L.; Langston, L.; Maupin, R.

submitted to the EMBL Data Library, February 1999

A:Description: The sequence of C. elegans cosmid H23N18.

A:Reference number: Z21449

A:Accession: T33983

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-141 <COU>

A:Cross-references: UNIPROT:O9TX27; EMBL:AF125959; PIDN:AA014734.1; GSPDB:GN00023; CESP:H

A:Experimental source: strain Bristol N2; clone H23N18

C:Gene: CESP:H23N18.5

A:Map position: 5

A:Introns: 33/2; 80/1; 121/3

Query Match 40.1%; Score 46.5; DB 2; Length 141;

Best Local Similarity 34.8%; Pred. No. 37;

Matches 8; Conservative 3; Mismatches 3; Indels 9; Gaps 1;

Qy 4 RCITRGFCRC-----ICIT 17

Db 99 RNVCDGFCCKAKNAKSLPLCT 121

RESULT 36

T49781

related to mutanase (mutA) gene [imported] - Neurospora crassa

N:Alternate names: protein B9J10.180

C:Species: Neurospora crassa

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004

C:Accession: T49781

R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022

A:Accession: T49781

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-883 <SCH>

A:Cross-references: UNIPROT:Q9P563; EMBL:AL356324; GSPDB:GN00116; NCSP:B9J10.180

A:Experimental source: BAC clone B9J10; strain OR74A

Search completed: October 26, 2004, 15:40:19
Job time : 33.5 secs

Search completed: October 26, 2004, 15:40:19
Job time : 33.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 15:13:05 ; Search time 115 Seconds
(without alignments)
56.149 Million cell updates/sec

Title: US-10-009-317a-32
Perfect score: 116
Sequence: 1 GFCRCICTRGFCRCICTR 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 202273 seqs, 35979299 residues

Total number of hits satisfying chosen parameters: 202273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 116 | 100.0 | 18 | 4 AAB35046 | Aab35046 Theta def |
| 2 | 116 | 100.0 | 18 | 5 ABP53295 | Abp53295 Anti-vira |
| 3 | 116 | 100.0 | 18 | 8 ADO35230 | Ado35230 Rhesus th |
| 4 | 111 | 95.7 | 18 | 8 ADO35240 | Ado35240 Rhesus th |
| 5 | 103 | 88.8 | 18 | 5 ABP53299 | Abp53299 Anti-vira |
| 6 | 101 | 87.1 | 18 | 4 AAB35030 | Aab35030 Theta def |
| 7 | 101 | 87.1 | 18 | 5 ABP53297 | Abp53297 Anti-vira |
| 8 | 101 | 87.1 | 18 | 6 AAE33866 | Aae33866 Macaca mu |
| 9 | 101 | 87.1 | 18 | 7 ADD95202 | Add95202 Cyclic de |
| 10 | 101 | 87.1 | 18 | 8 ADJ35357 | Adj35357 Antimicro |
| 11 | 101 | 87.1 | 18 | 8 ADG70012 | Adg70012 Rhesus th |
| 12 | 101 | 87.1 | 18 | 8 ADO35229 | Ado35229 Rhesus th |
| 13 | 101 | 87.1 | 18 | 8 ADO35228 | Ado35228 Rhesus th |
| 14 | 101 | 87.1 | 18 | 8 ADO35239 | Ado35239 Rhesus th |
| 15 | 101 | 87.1 | 18 | 8 ADO35250 | Ado35250 Rhesus th |
| 16 | 101 | 87.1 | 38 | 8 ADO35263 | Ado35263 Monkey RT |
| 17 | 100 | 86.2 | 18 | 8 ADO35255 | Ado35255 Rhesus th |
| 18 | 93 | 80.2 | 18 | 6 AAE33805 | Aae33805 I1Y retro |
| 19 | 93 | 80.2 | 18 | 6 AAE33806 | Aae33806 I1Y retr |
| 20 | 93 | 80.2 | 18 | 8 ADN08180 | Adn08180 Human ret |
| 21 | 93 | 80.2 | 18 | 8 ADN08181 | Adn08181 Human ret |
| 22 | 91 | 78.4 | 18 | 4 AAB35037 | Aab35037 Rhesus ma |
| 23 | 90 | 77.6 | 18 | 5 ABP53294 | Abp53294 Synthetic |
| 24 | 90 | 77.6 | 18 | 6 AAE33801 | Aae33801 Human ret |
| 25 | 90 | 77.6 | 18 | 6 AAE33863 | Aae33863 Enantio-r |

| | | | | | |
|----|----|------|----|------------|--------------------|
| 26 | 90 | 77.6 | 18 | 8 ADN08176 | Adn08176 Human ret |
| 27 | 89 | 76.7 | 18 | 8 ADO35249 | Ado35249 Rhesus th |
| 28 | 88 | 75.9 | 18 | 5 ABP53298 | Abp53298 Anti-vira |
| 29 | 87 | 75.0 | 18 | 5 AAU91017 | Aau91017 Transplan |
| 30 | 87 | 75.0 | 18 | 6 AAE33864 | Aae33864 Enantio-R |
| 31 | 87 | 75.0 | 18 | 6 AAE33802 | Aae33802 R9K retro |
| 32 | 87 | 75.0 | 18 | 8 ADN08177 | Adn08177 Human ret |
| 33 | 86 | 74.1 | 18 | 4 AAB35047 | Aab35047 Theta def |
| 34 | 86 | 74.1 | 18 | 5 ABP53296 | Abp53296 Anti-vira |
| 35 | 86 | 74.1 | 18 | 8 ADO35231 | Ado35231 Rhesus th |
| 36 | 86 | 74.1 | 18 | 8 ADO35242 | Ado35242 Rhesus th |
| 37 | 85 | 74.1 | 18 | 8 ADO35241 | Ado35241 Rhesus th |
| 38 | 85 | 73.3 | 18 | 6 AAE33804 | Aae33804 I15Y retr |
| 39 | 85 | 73.3 | 18 | 6 AAE33803 | Aae33803 I6Y retro |
| 40 | 85 | 73.3 | 18 | 8 ADO35256 | Ado35256 Rhesus th |
| 41 | 85 | 73.3 | 18 | 8 ADN08179 | Adn08179 Human ret |
| 42 | 85 | 73.3 | 18 | 8 ADN08178 | Adn08178 Human ret |
| 43 | 83 | 71.6 | 18 | 6 AAE33807 | Aae33807 R4Y retro |
| 44 | 83 | 71.6 | 18 | 8 ADO35246 | Ado35246 Rhesus th |
| 45 | 83 | 71.6 | 18 | 8 ADO35245 | Ado35245 Rhesus th |

ALIGNMENTS

RESULT 1
AAB35046
ID AAB35046 standard; peptide; 18 AA.
XX
AC AAB35046;
XX
DT 27-MAR-2001 (first entry)
XX
DE Theta defensin SEQ ID NO: 30.
XX
KW Theta defensin; antimicrobial; cyclic; bacterium; fungus; protozoan;
KW virus; helminth; disinfectant; food preservative; analogue.
XX
OS Unidentified.
XX
PN WO200068265-A1.
XX
PD 16-NOV-2000.
XX
PF 10-MAY-2000; 2000WO-US012842.
XX
PR 10-MAY-1999; 99US-00309487.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Selssted ME, Tang Y, Yuan J, Ouellette AJ;
XX
DR WPI; 2001-031853/04.
XX
PT Novel theta defensin peptide with antimicrobial activity against
PT bacteria, yeast, fungi, protozoa and viruses.
XX
PS Claim 15; Fig 16; 110pp; English.
XX
CC The present invention provides theta defensin peptides and analogues
CC which have antimicrobial activity. They can be used in the treatment of
CC bacterial, viral, fungal, protozoan and helminthic infections, in
CC disinfectants and as food preservatives
XX
SQ Sequence 18 AA;
Query Match 100.0%; Score 116; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 8.1e-06; Gaps 0;
Matches 18; Conservative 0; Mismatches 0; Indels 0;

OY 1 GFCRCICTRGFCRCICTR 18
DB 1 GFCRCICTRGFCRCICTR 18

RESULT 2
 ID ABP53295 standard; peptide; 18 AA.
 AC ABP53295;
 DT 13-NOV-2002 (first entry)
 DE Anti-viral theta defensin peptide RTD-2 SEQ ID NO:28.
 KW Anti-viral; viral infection; theta-defensin; lipid environment;
 KW amphipathic alpha-helical structure; virucide; anti-HIV; immunisation;
 KW viral growth inhibitor; viral proliferation inhibitor.
 XX
 OS Macaca mulatta.
 OS Synthetic.
 PN WO200260468-A2.
 XX
 PD 08-AUG-2002.
 XX
 XX 29-JAN-2002; 2002WO-US002435.
 XX
 PR 30-JAN-2001; 2001US-0265270P.
 PR 01-AUG-2001; 2001US-0309368P.
 XX
 XX (IOWA) UNIV IOWA RES FOUND.
 XX
 XX Maury W, Stapleton J, Stinski M, Roller R, Mccray PB, Tack B;
 XX WPI; 2002-674815/72.
 XX
 PT New method of using a first anti-viral peptide comprising a Theta-
 PT defensin peptide in an amphipathic Alpha-helical structure in a lipid
 PT environment for reducing the infectivity of a virus.
 XX
 PS Disclosure; Page 10; 65pp; English.
 XX
 CC The present invention describes a method (M1) of using a first anti-viral
 CC peptide (I) comprising a theta-defensin peptide in an amphipathic alpha-
 CC helical structure in a lipid environment for reducing the infectivity of
 CC a virus. (I) can have virucide and anti-HIV activities, and can be used
 CC to reduce virus growth, infectivity burden, shed, and development of anti
 CC -viral resistance. (I) can be used for inhibiting the growth and
 CC proliferation of a virus and so can be used for: (a) protecting or
 CC treating subject from a viral infection, preventing recurrent viral
 CC infection in a subject harbouring a latent virus, controlling virus
 CC spread within a virally-infected subject (VS), reducing viral burden in a
 CC VS, reducing virus shed from a VS, reducing percentage of VS in a
 CC population regardless of viral infection status, or inducing latency in a
 CC VS; (b) reducing the infectivity of a virus; and (c) rendering virus-
 CC contaminated tissue or fluid sample safe for use, or reducing the number
 CC of infectious virus particles in a population of viruses. (M1) is useful
 CC for reducing the infectivity of a virus in sheep, cattle, horses, swine,
 CC cats, fowl and humans e.g. an enveloped virus infecting humans such as
 CC human immunodeficiency virus (HIV). Preferably, the anti-viral peptide is
 CC administered to a patient who is immunosuppressed or to a subject who is
 CC not infected with the virus, where the first anti-viral peptide is
 CC administered prior to or subsequent to the virus contacting the subject.
 CC The anti-viral peptide is most preferably administered to a subject who
 CC is chronically, latently or acutely infected with the virus. The present
 CC sequence represents a rhesus monkey theta defensin anti-viral peptide,
 CC which is given in the exemplification of the present invention
 XX
 SQ Sequence 18 AA;
 Query Match 100.0%; Score 116; DB 5; Length 18;
 Best Local Similarity 100.0%; Pred. No. 8.1e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GFCRCICTRGFCRCICTR 18

Db
 RESULT 3
 ID ADO35230 standard; peptide; 18 AA.
 AC ADO35230;
 DT 15-JUL-2004 (first entry)
 DE Rhesus theta defensin peptide, RTD-2.
 KW Monkey; Rhesus theta defensin; RTD-2; antimicrobial peptide; cyclic;
 KW antimicrobial; antiinflammatory; antibacterial; virucide; fungicide;
 KW food; contact lens solution; eye wash solution; inflammatory response;
 KW microbicidal inhibition; microbiostatic growth inhibition; disinfectant;
 KW food preservative; bacterial infection; viral infection;
 KW fungal infection; haemolytic activity.
 XX
 OS Macaca mulatta.
 XX
 XX Key Location/Qualifiers
 XX Modified-site 1. .18
 FT /note= "The peptide is cyclised by a covalent link
 FT between these two residues"
 FT Disulfide-bond 3. .16
 FT Disulfide-bond 5. .14
 FT Disulfide-bond 7. .12
 XX
 PN US2004014669-A1.
 XX
 XX 22-JAN-2004.
 PD
 XX 30-APR-2003; 2003US-00427715.
 PF
 XX 30-APR-2002; 2002US-0377071P.
 PR
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX Selsted ME, Tran DQ;
 XX WPI; 2004-167945/16.
 DR
 XX Novel theta defensin analog useful for reducing or inhibiting growth or
 PT survival of a microorganism in an environment such as food or food
 PT product, contact lens solution, or eye wash solution, an inanimate
 PT object.
 XX
 XX Example 1; SEQ ID NO 2; 46pp; English.
 XX
 CC The invention relates to a theta defensin analogue defined by formulae
 CC detailed in the claims or appearing as ADO35239-ADO35257. The theta
 CC defensin analogue is useful for reducing or inhibiting growth or survival
 CC of a microorganism in an environment capable of sustaining the growth or
 CC survival of the microorganism and is useful for reducing or inhibiting
 CC growth or survival of a microorganism in an environment such as food or
 CC food product, a solution (e.g., contact lens solution, or eye wash
 CC solution), an inanimate object comprising surface, or a mammal. The
 CC peptides are also useful for decreasing inflammatory response and for
 CC microbicidal inhibition of survival of microorganism as well as
 CC microbistatic inhibition of growth. Thus the peptides are useful as
 CC therapeutic agents, disinfectants, food preservatives, or medicaments.
 CC The peptides are also useful for treating a patient suffering from
 CC bacterial, viral, fungal or other infection. The theta defensins have
 CC high antimicrobial activity and low haemolytic activity. The present
 CC sequence represents the rhesus monkey wild-type theta defensin RTD-2.
 XX
 SQ Sequence 18 AA;
 Query Match 100.0%; Score 116; DB 8; Length 18;
 Best Local Similarity 100.0%; Pred. No. 8.1e-06;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
|||||
Db 1 GFCRCICTRGFCRCICTR 18
|||||

RESULT 4
ADO35240
ID ADO35240 standard; peptide; 18 AA.
AC ADO35240;
DT 15-JUL-2004 (first entry)
DE Rhesus theta defensin analogue peptide arTD-2-OH.
XX
KW Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial;
antiflammatory; antibacterial; virucide; fungicide; food;
contact lens solution; eye wash solution; inflammatory response;
KW microbicidal inhibition; microbistatic growth inhibition; disinfectant;
KW food preservative; bacterial infection; viral infection;
KW fungal infection; haemolytic activity.
XX
OS Macaca mulatta.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 3..16
FT Disulfide-bond 5..14
FT Disulfide-bond 7..12
FT Modified-site 18 /note= "Hydroxylated"
FT
XX US2004014669-A1.
XX
XX 22-JAN-2004.
XX
XX 30-APR-2003; 2003US-00427115.
XX
XX 30-APR-2002; 2002US-0377071P.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Selsted ME, Tran DQ;
XX WPI; 2004-167945/16.
XX
XX Novel theta defensin analog useful for reducing or inhibiting growth or
survival of a microorganism in an environment such as food or food
product, contact lens solution, or eye wash solution, an inanimate
object.
XX
XX Claim 1; SEQ ID NO 14; 46pp; English.
XX
XX The invention relates to a theta defensin analogue defined by formulae
detailed in the claims or appearing as ADO35239-ADO35257. The theta
defensin analogue is useful for reducing or inhibiting growth or survival
of a microorganism in an environment capable of sustaining the growth or
survival of the microorganism and is useful for reducing or inhibiting
growth or survival of a microorganism in an environment such as food or
food product, a solution (e.g. contact lens solution, or eye wash
solution), an inanimate object comprising surface, or a mammal. The
peptides are also useful for decreasing inflammatory response and for
microbicidal inhibition of survival of microorganism as well as
microbistatic inhibition of growth. Thus the peptides are useful as
therapeutic agents, disinfectants, food preservatives, or medicaments.
XX The peptides are also useful for treating a patient suffering from
bacterial, viral, fungal or other infection. The theta defensins have
high antimicrobial activity and low haemolytic activity. The present
sequence represents a Rhesus theta defensin analogue peptide.
XX
SQ Sequence 18 AA;

Query Match 95.7%; Score 111; DB 8; Length 18;
Best Local Similarity 94.4%; Pred. No. 2.7e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
|||||
Db 1 GFCRCICTRGFCRCICTR 18
|||||

RESULT 5
ABP53299
ID ABP53299 standard; peptide; 18 AA.
XX
AC ABP53299;
DT 13-NOV-2002 (first entry)
DE Anti-viral chimeric theta defensin peptide H/RTD-2 SEQ ID NO:32.
XX
KW Anti-viral; viral infection; theta-defensin; lipid environment;
amphipathic alpha-helical structure; virucide; anti-HIV; immunisation;
KW viral growth inhibitor; viral proliferation inhibitor.
XX
OS Homo sapiens.
OS Macaca mulatta.
OS Synthetic.
XX WO200260468-A2.
XX
XX 08-AUG-2002.
XX
XX 29-JAN-2002; 2002WO-US0002435.
XX
XX 30-JAN-2001; 2001US-0265270P.
XX
XX 01-AUG-2001; 2001US-0309368P.
XX
XX (IOWA) UNIV IOWA RES FOUND.
XX
XX Maury W, Stapleton J, Stinski M, Roller R, Mcray PB, Tack B;
XX WPI; 2002-674815/72.
XX
XX New method of using a first anti-viral peptide comprising a Theta-
defensin peptide in an amphipathic Alpha-helical structure in a lipid
environment for reducing the infectivity of a virus.
XX
XX Disclosure; Page 10; 65pp; English.
XX
XX The present invention describes a method (M1) of using a first anti-viral
peptide (I) comprising a theta-defensin peptide in an amphipathic alpha-
helical structure in a lipid environment for reducing the infectivity of
a virus. (I) can have virucide and anti-HIV activities, and can be used
to reduce virus growth, infectivity burden, shed, and development of anti-
viral resistance. (I) can be used for inhibiting the growth and
proliferation of a virus and so can be used for; (a) protecting or
treating subject from a viral infection, preventing recurring viral
infection in a subject harbouring a latent virus, controlling virus
spread within a virally-infected subject (VS), reducing viral burden in a
VS, reducing virus shed from a VS, reducing percentage of VS in a
population regardless of viral infection status, or inducing latency in a
VS; (b) reducing the infectivity of a virus; and (c) rendering virus-
contaminated tissue or fluid sample safe for use, or reducing the number
of infectious virus particles in a population of viruses. (M1) is useful
for reducing the infectivity of a virus in sheep, cattle, horses, swine,
cats, fowl and humans e.g. an enveloped virus infecting humans such as
human immunodeficiency virus (HIV). Preferably, the anti-viral peptide is
administered to a patient who is immunosuppressed or to a subject who is
not infected with the virus, where the first anti-viral peptide is
administered prior to or subsequent to the virus contacting the subject.
XX The anti-viral peptide is most preferably administered to a subject who
is chronically, latently or acutely infected with the virus. The present
sequence represents a chimeric human/rhesus monkey theta defensin anti-

RESULT 8
AAE33866
ID AAE33866 standard; peptide; 18 AA.

AC AAE33866;
XX
DT 16-APR-2003 (first entry)

DE Macaca mulatta RTD1 peptide.

XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
KW antibiotic modelling; antimicrobial; rhesus monkey; theta defensin 1A;
KW RTD1.

XX Macaca mulatta.

XX WO200285401-A1.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012353.

XX 18-APR-2001; 2001US-0284855P.

XX (REGC) UNIV CALIFORNIA.

XX Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2003-103387/99.

XX New isolated retrocyclin peptide, useful for preventing retroviral
PT infections in cells susceptible to bacterial or viral infections or
PT treating patients having the infections, such as HIV, sexually
PT transmitted diseases, vaginosis.

XX Example 1; Fig 3C; 72pp; English.

XX The invention relates to novel retrocyclin peptides. Peptides and methods
CC of the invention are useful for preventing retroviral infections in cells
CC susceptible to bacterial or viral infections, or treating patients having
CC infections such as HIV (human immunodeficiency virus), sexually
CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
CC retrocyclin-mediated killing is useful for modelling and screening novel
CC antibiotics. The invention is also useful in gene therapy. The present
CC sequence is rhesus monkey theta defensin, RTD1 peptide. This sequence is
CC used in the exemplification of the invention

XX Sequence 18 AA;

Query Match 87.1%; Score 101; DB 6; Length 18;

Best Local Similarity 83.3%; Pred. No. 0.0003;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
DB 1 GVCRCICTRGFCRCICLR 18

RESULT 9
ADD95202

ID ADD95202 standard; peptide; 18 AA.

XX ADD95202;

XX 29-JAN-2004 (first entry)

XX Cyclic defensin fragment.

XX bacterial infection; human pathogen; holin; defensin;

XX peptide nucleic acid; PNA; penicillin; tetracycline; ampicillin;

XX kanamycin; antibiotic; antibacterial; antibiotic-resistance gene; cyclic.

OS Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 1. .18
FT /note= "Residue 1 and residue 18 bond to form a cyclic moiety;"

FT Disulfide-bond 3. .16

FT Disulfide-bond 5. .14

XX WO2003059392-A2.

XX 24-JUL-2003.

XX 17-JAN-2003; 2003WO-DE000124.

XX 18-JAN-2002; 2002DE-01001862.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Braun K, Braun I, Debus J, Pipkorn R, Waldeck W;

XX WPI; 2003-689464/65.

XX New conjugate of transport mediator and active agent, useful for treating
XX prokaryotic infections, especially by neutralizing antibiotic resistance
XX gene.

XX Disclosure; Fig 10; 34pp; German.

XX This invention describes a novel conjugate for treating prokaryotic
CC infections which comprises a transport mediator for passage through the
CC prokaryotic cell membrane and a compound, directed against a prokaryote
CC and intended for introduction into it. The prokaryote is a bacterium,
CC especially one pathogenic in humans. The transport mediator is preferably
CC a human peptide or protein, especially a phage-holin protein, its active
CC fragment or variant or a defensin. The introduced compound is a peptide
CC nucleic acid (PNA) that inhibits a gene, especially one implicated in
CC resistance to penicillin, tetracycline, ampicillin or kanamycin. The
CC conjugate has the structure transport mediator-spacer-PNA where the
CC spacer is poly(glycine and/or lysine), preferably containing 2-6 amino
CC acids and the spacer is linked to the transport mediator through a
CC cleavable disulfide bridge. The conjugates are administered together with
CC an antibiotic, by parenteral, transdermal or subcutaneous routes. The
CC products of the invention have antibacterial activity and are used,
CC especially in combination with antibiotics, for treating prokaryotic,
CC specifically bacterial, infections, especially where the pathogen is
CC resistant to at least one antibiotic and then the PNA is directed against
CC the antibiotic-resistance gene. Where the PNA is directed against an
CC antibiotic resistance gene the conjugate will render the bacteria
CC sensitive to co-administered antibiotics i.e. 'old' antibiotics can be
CC used successfully in cases where normally they would be ineffective. This
CC sequence represents a cyclic defensin fragment described in the
CC disclosure of the invention.

XX Sequence 18 AA;

Query Match 87.1%; Score 101; DB 7; Length 18;

Best Local Similarity 83.3%; Pred. No. 0.0003;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18

DB 1 GFCRCICLRGVCRCICTR 18

RESULT 10

ADD35357

ID ADD35357 standard; peptide; 18 AA.

XX ADD35357;

XX 15-JAN-2004 (first entry)

XX Antimicrobial peptide theta-defensin.
 DE
 XX antimicrobial; ophthalmic; prostaglandin; hypotensive; ophthalmological;
 XX intraocular pressure; glaucoma; ocular hypertension; hyperaemia;
 XX irritation; inflammation; conjunctiva; ocular cell dysplasia;
 XX iridial melanocyte hyperplasia; hyperpigmentation.
 XX Unidentified.
 OS
 XX WO2003079997-A2.
 PN
 XX 02-OCT-2003.
 PD
 XX 21-MAR-2003; 2003WO-US008935.
 PF
 XX 21-MAR-2002; 2002US-0367071P.
 PR
 XX (CAYM-) CAYMAN CHEM CO.
 XX
 XX Maxey KM, Johnson J;
 PA
 XX WPI; 2004-011506/01.
 DR
 XX Ophthalmic solution useful for the treatment of increased intraocular
 XX pressure comprises a prostaglandin of the F-series and an antimicrobial
 XX peptide.
 PT
 XX Disclosure; Page 11; 11pp; English.
 PS
 XX The invention relates to a novel ophthalmic solution comprising a
 XX prostaglandin of the F-series and an antimicrobial peptide. A solution of
 XX the invention has hypotensive and ophthalmological activity. The solution
 XX is useful for the treatment of increased intraocular pressure, such as
 XX caused by glaucoma and for the reduction of ocular hypertension. The
 XX prostaglandin and the antimicrobial peptide work synergistically to
 XX provide beneficial reduction in the incidence of irritant and toxic side
 XX effects such as hyperaemia, irritation and inflammation of conjunctiva,
 XX ocular cell dysplasia, iridial melanocyte hyperplasia, and
 XX hyperpigmentation, associated with the prior art prostaglandin
 XX compositions. The present sequence represents an antimicrobial peptide of
 XX the invention.
 CC
 XX Sequence 18 AA;
 SQ

Query Match 87.1%; Score 101; DB 8; Length 18;
 Best Local Similarity 83.3%; Pred. No. 0.0003;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GFCRCICTRGFCRCICTR 18
 |||||:|||||
 Db 1 GFCRCICRRGVRCICTR 18

RESULT 11
 ADG70012
 ID ADG70012 standard; peptide; 18 AA.
 XX
 AC ADG70012;
 XX
 DT 11-MAR-2004 (first entry)
 DE Rhesus theta-defensin-1 (RTD-1) peptide.
 XX
 XX rhesus theta defensin-1; RTD-1; bacteraemia; lipopolysaccharide; LPS;
 XX lipoteichoic acid; LTA; septic shock; antibacterial; fungicide; virucide;
 XX immunomodulator; anticoagulant activity;
 XX microbial cell-wall biosynthesis; immunomodulation; anticoagulant.
 OS
 XX Macaca mulatta.
 PN
 XX WO2003105883-A1.
 XX

PD 24-DEC-2003.
 XX
 XX 30-MAY-2003; 2003WO-EP005694.
 PF
 XX 13-JUN-2002; 2002DE-01026216.
 PR
 XX (FARB) BAYER HEALTHCARE AG.
 PA
 XX Ladel C, Newton B, Labischinski H, Brunner N, Gerdes C;
 PI
 XX WPI; 2004-071500/07.
 DR
 XX Use of rhesus theta defensin-1 for treating or preventing bacteremia and
 XX septic shock, also for binding bacterial products and as immunomodulator
 XX and anticoagulant.
 PT
 XX Example 1; SEQ ID NO 1; 28pp; German.
 PS
 XX This invention describes the novel use of rhesus theta defensin-1 (RTD-1)
 XX for preparing a composition for treatment and/or prevention of
 XX bacteraemia for binding bacterial products such as lipopolysaccharide
 XX (LPS) and/or lipoteichoic acid (LTA), or for treatment and/or prevention
 XX of septic shock. RTD-1, isolated from immune cells of rhesus monkeys, has
 XX antibacterial, fungicide, virucide, immunomodulator and anticoagulant
 XX activity. RTD-1 inhibits microbial cell-wall biosynthesis and also binds
 XX to LPS and LTA. RTD-1 is useful for treatment and prevention of severe
 XX infections caused by Gram-positive or -negative bacteria and yeasts, or
 XX by viruses. RTD-1 combines four advantageous properties: a direct
 XX antimicrobial action, neutralisation of bacterial products (by binding),
 XX immunomodulation (reducing release of proinflammatory cytokines but
 XX increasing release of regulatory factors) and anticoagulant action, so
 XX provides a better and simpler treatment.
 CC
 XX Sequence 18 AA;
 SQ

Query Match 87.1%; Score 101; DB 8; Length 18;
 Best Local Similarity 83.3%; Pred. No. 0.0003;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GFCRCICTRGFCRCICTR 18
 |||||:|||||
 Db 1 GFCRCICRRGVRCICTR 18

RESULT 12
 ADO35229
 ID ADO35229 standard; peptide; 18 AA.
 XX
 AC ADO35229;
 XX
 DT 15-JUL-2004 (first entry)
 DE Rhesus theta defensin peptide, RTD-1.
 XX
 XX Monkey; Rhesus theta defensin; RTD-1; antimicrobial peptide; cyclic;
 XX antimicrobial; antiinflammatory; antibacterial; virucide; fungicide;
 XX food; contact lens solution; eye wash solution; inflammatory response;
 XX microbicidal inhibition; microbistatic growth inhibition; disinfectant;
 XX food preservative; bacterial infection; viral infection;
 XX fungal infection; haemolytic activity.
 OS
 XX Macaca mulatta.
 PF
 XX Key Location/Qualifiers
 XX Modified-site 1..18
 XX /note= "the peptide is cyclised by a covalent link
 XX between these two residues"
 XX Disulfide-bond 3..16
 XX Disulfide-bond 5..14
 XX Disulfide-bond 7..12
 XX
 PN US2004014669-A1.
 XX

```

PD 22-JAN-2004.
XX
XX
XX 30-APR-2003; 2003US-00427715.
XX
XX 30-APR-2002; 2002US-0377071P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Selsted ME, Tran DQ;
XX
XX WPI; 2004-167945/16.
XX
XX Novel theta defensin analog useful for reducing or inhibiting growth or
PT survival of a microorganism in an environment such as food or food
PT product, contact lens solution, or eye wash solution, an inanimate
PT object.
XX
XX Example 1; SEQ ID NO 1; 46pp; English.
XX
XX The invention relates to a theta defensin analogue defined by formulae
CC detailed in the claims or appearing as ADO35239-ADO35257. The theta
CC defensin analogue is useful for reducing or inhibiting growth or survival
CC of a microorganism in an environment capable of sustaining the growth or
CC survival of the microorganism and is useful for reducing or inhibiting
CC growth or survival of a microorganism in an environment such as food or
CC food product, a solution (e.g., contact lens solution, or eye wash
CC solution), an inanimate object comprising surface, or a mammal. The
CC peptides are also useful for decreasing inflammatory response and for
CC microbicidal inhibition of survival of microorganism as well as
CC therapeutic agents, disinfectants, food preservatives, or medicaments.
CC The peptides are also useful for treating a patient suffering from
CC bacterial, viral, fungal or other infection. The theta defensins have
CC high antimicrobial activity and low haemolytic activity. The present
CC sequence represents the rhesus monkey wild-type theta defensin RTD-1.
XX
XX Sequence 18 AA;
XX
XX Query Match 87.1%; Score 101; DB 8; Length 18;
XX Best Local Similarity 83.3%; Pred. No. 0.0003;
XX Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 GFCRCICTRGFCRCICTR 18
XX |||||:|||||
XX Db 1 GFCRCICRGVCRCICTR 18
XX |||||:|||||
XX
XX RESULT 13
XX ADO35238
XX ID ADO35238 standard; peptide; 18 AA.
XX
XX AC ADO35238;
XX
XX DT 15-JUL-2004 (first entry)
XX
XX DE Rhesus theta defensin analogue peptide aRTD-1-OH.
XX
XX KW Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial;
XX antiinflammatory; antibacterial; virucide; fungicide; food;
XX contact lens solution; eye wash solution; inflammatory response;
XX microbicidal inhibition; microbistatic growth inhibition; disinfectant;
XX food preservative; bacterial infection; viral infection;
XX fungal infection; haemolytic activity.
XX
XX OS Macaca mulatta.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Disulfide-bond 3..16
XX FT Disulfide-bond 5..14
XX FT Disulfide-bond 7..12
XX FT Modified-site 18
XX FT /note= "Hydroxylated"
XX
XX US2004014669-A1.
XX
XX 22-JAN-2004.
XX
XX 30-APR-2003; 2003US-00427715.
XX
XX 30-APR-2002; 2002US-0377071P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Selsted ME, Tran DQ;
XX
XX WPI; 2004-167945/16.
XX
XX Novel theta defensin analog useful for reducing or inhibiting growth or
PT survival of a microorganism in an environment such as food or food
PT product, contact lens solution, or eye wash solution, an inanimate
PT object.
XX
XX Example 2; SEQ ID NO 12; 46pp; English.
XX
XX The invention relates to a theta defensin analogue defined by formulae
CC detailed in the claims or appearing as ADO35239-ADO35257. The theta
CC defensin analogue is useful for reducing or inhibiting growth or survival
CC of a microorganism in an environment capable of sustaining the growth or
CC survival of the microorganism and is useful for reducing or inhibiting
CC growth or survival of a microorganism in an environment such as food or
CC food product, a solution (e.g., contact lens solution, or eye wash
CC solution), an inanimate object comprising surface, or a mammal. The
CC peptides are also useful for decreasing inflammatory response and for
CC microbicidal inhibition of survival of microorganism as well as
CC therapeutic agents, disinfectants, food preservatives, or medicaments.
CC The peptides are also useful for treating a patient suffering from
CC bacterial, viral, fungal or other infection. The theta defensins have
CC high antimicrobial activity and low haemolytic activity. The present
CC sequence represents a Rhesus theta defensin analogue peptide.
XX
XX Sequence 18 AA;
XX
XX Query Match 87.1%; Score 101; DB 8; Length 18;
XX Best Local Similarity 83.3%; Pred. No. 0.0003;
XX Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 GFCRCICTRGFCRCICTR 18
XX |||||:|||||
XX Db 1 GFCRCICRGVCRCICTR 18
XX |||||:|||||
XX
XX RESULT 14
XX ADO35239
XX ID ADO35239 standard; peptide; 18 AA.
XX
XX AC ADO35239;
XX
XX DT 15-JUL-2004 (first entry)
XX
XX DE Rhesus theta defensin analogue peptide aRTD-1-NH.
XX
XX KW Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial;
XX antiinflammatory; antibacterial; virucide; fungicide; food;
XX contact lens solution; eye wash solution; inflammatory response;
XX microbicidal inhibition; microbistatic growth inhibition; disinfectant;
XX food preservative; bacterial infection; viral infection;
XX fungal infection; haemolytic activity.
XX
XX OS Macaca mulatta.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Disulfide-bond 3..16
XX FT Disulfide-bond 5..14
XX FT Disulfide-bond 7..12
XX FT Modified-site 18
XX FT /note= "Hydroxylated"
XX

```

FT Disulfide-bond 7. 12
 FT Modified-site 18
 FT /note= "Amidated"
 XX
 FN US2004014669-A1.
 XX
 XX 22-JAN-2004.
 XX
 XX 30-APR-2003; 2003US-00427715.
 XX
 PR 30-APR-2002; 2002US-0377071P.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX Selsted ME, Tran DQ;
 XX
 XX WPI; 2004-167945/16.
 XX
 XX Novel theta defensin analog useful for reducing or inhibiting growth or survival of a microorganism in an environment such as food or food product, contact lens solution, or eye wash solution, an inanimate object.
 XX
 XX Claim 1; SEQ ID NO 13; 46pp; English.
 XX
 CC The invention relates to a theta defensin analogue defined by formulae detailed in the claims or appearing as ADO35239-ADO35257. The theta defensin analogue is useful for reducing or inhibiting growth or survival of a microorganism in an environment capable of sustaining the growth or survival of the microorganism and is useful for reducing or inhibiting growth or survival of a microorganism in an environment such as food or food product, a solution (e.g., contact lens solution, or eye wash solution), an inanimate object comprising surface, or a mammal. The peptides are also useful for decreasing inflammatory response and for microbicidal inhibition of survival of microorganism as well as therapeutic agents, disinfectants, food preservatives, or medicaments. The peptides are also useful for treating a patient suffering from bacterial, viral, fungal or other infection. The theta defensins have high antimicrobial activity and low haemolytic activity. The present sequence represents a Rhesus theta defensin analogue peptide.
 XX
 SQ Sequence 18 AA;
 Query Match 87.1%; Score 101; DB 8; Length 18;
 Best Local Similarity 83.3%; Pred. No. 0.0003;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GFCRCICTRGFCRCICTR 18
 Db 1 GFCRCICTRGFCRCICTR 18
 RESULT 15
 ADO35250
 ID ADO35250 standard; peptide; 18 AA.
 XX
 AC ADO35250;
 XX
 XX 15-JUL-2004 (first entry)
 DT
 XX Rhesus theta defensin analogue peptide RTD-1-26.
 DE
 XX Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial; antiinflammatory; antibacterial; virucide; fungicide; food; contact lens solution; eye wash solution; inflammatory response; microbicidal inhibition; microbicidal growth inhibition; disinfectant; food preservative; bacterial infection; viral infection; fungal infection; haemolytic activity; cyclic.
 XX
 OS Macaca mulatta.
 OS Synthetic.

PH Key Location/Qualifiers
 FT Modified-site 1. 18
 FT /note= "The peptide is cyclised by a covalent link between these two residues"
 FT Disulfide-bond 3. 16
 FT Disulfide-bond 5. 14
 FT Disulfide-bond 7. 12
 XX
 PN US2004014669-A1.
 XX
 XX 22-JAN-2004.
 PD
 XX 30-APR-2003; 2003US-00427715.
 XX
 PR 30-APR-2002; 2002US-0377071P.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX Selsted ME, Tran DQ;
 XX
 XX WPI; 2004-167945/16.
 XX
 XX Novel theta defensin analog useful for reducing or inhibiting growth or survival of a microorganism in an environment such as food or food product, contact lens solution, or eye wash solution, an inanimate object.
 XX
 XX Claim 1; SEQ ID NO 24; 46pp; English.
 XX
 CC The invention relates to a theta defensin analogue defined by formulae detailed in the claims or appearing as ADO35239-ADO35257. The theta defensin analogue is useful for reducing or inhibiting growth or survival of a microorganism in an environment capable of sustaining the growth or survival of the microorganism and is useful for reducing or inhibiting growth or survival of a microorganism in an environment such as food or food product, a solution (e.g., contact lens solution, or eye wash solution), an inanimate object comprising surface, or a mammal. The peptides are also useful for decreasing inflammatory response and for microbicidal inhibition of survival of microorganism as well as therapeutic agents, disinfectants, food preservatives, or medicaments. The peptides are also useful for treating a patient suffering from bacterial, viral, fungal or other infection. The theta defensins have high antimicrobial activity and low haemolytic activity. The present sequence represents a Rhesus theta defensin analogue peptide.
 XX
 SQ Sequence 18 AA;
 Query Match 87.1%; Score 101; DB 8; Length 18;
 Best Local Similarity 88.9%; Pred. No. 0.0003;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GFCRCICTRGFCRCICTR 18
 Db 1 GFCRCICTRGFCRCICTR 18
 RESULT 16
 ADO35263
 ID ADO35263 standard; protein; 38 AA.
 XX
 AC ADO35263;
 XX
 XX 15-JUL-2004 (first entry)
 DT
 XX Monkey RTD-1 (rhesus theta defensin 1) 2X protein.
 DE
 XX Monkey; Rhesus theta defensin; RTD-1; antimicrobial peptide; antimicrobial; antiinflammatory; antibacterial; virucide; fungicide; food; contact lens solution; eye wash solution; inflammatory response; microbicidal inhibition; microbicidal growth inhibition; disinfectant; food preservative; bacterial infection; viral infection; fungal infection; haemolytic activity.

```

XX OS Macaca mulatta.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Cleavage-site 1..2 "CNBr cleavage site"
XX FT Peptide 2..19 /note= "RTD-1"
XX FT Cleavage-site 20..21 /note= "CNBr cleavage site"
XX FT Peptide 21..38 /note= "RTD-1"
XX US2004014669-A1.
XX PD 22-JAN-2004.
XX PF 30-APR-2003; 2003US-00427715.
XX PR 30-APR-2002; 2002US-0377071P.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Selsted ME, Tran DQ;
XX DR WPI; 2004-167945/16.
XX DR N-PSDB; ADO35262.
XX PT Novel theta defensin analog useful for reducing or inhibiting growth or
XX PT survival of a microorganism in an environment such as food or food
XX PT product, contact lens solution, or eye wash solution, an inanimate
XX PT Object.
XX PS Example 4; Fig 17a; 46pp; English.
XX CC The invention relates to a theta defensin analogue defined by formulae
XX CC detailed in the claims or appearing as ADO35239-ADO35257. The theta
XX CC defensin analogue is useful for reducing or inhibiting growth or survival
XX CC of a microorganism in an environment capable of sustaining the growth or
XX CC survival of the microorganism and is useful for reducing or inhibiting
XX CC growth or survival of a microorganism in an environment such as food or
XX CC food product, a solution (e.g., contact lens solution, or eye wash
XX CC solution), an inanimate object comprising surface, or a mammal. The
XX CC peptides are also useful for decreasing inflammatory response and for
XX CC microbistatic inhibition of survival of microorganism as well as
XX CC therapeutic agents, disinfectants, food preservatives, or medicaments.
XX CC The peptides are also useful for treating a patient suffering from
XX CC bacterial, viral, fungal or other infection. The theta defensins have
XX CC high antimicrobial activity and low haemolytic activity. The present
XX CC sequence is protein containing 2 copies of the rhesus monkey theta
XX CC defensin, RTD-1 peptide.
XX SQ Sequence 38 AA;
XX Query Match 87.1%; Score 101; DB 8; Length 38;
XX Best Local Similarity 83.3%; Pred. No. 0.00052;
XX Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX QY 1 GFCRCICTRGFCRCICTR 18
XX Db |||||:|||||
XX 2 GFCRCICRRGVCRCICTR 19
XX RESULT 17
XX ADO35255
XX ID ADO35255 standard; peptide; 18 AA.
XX AC ADO35255;
XX XX
XX DT 15-JUL-2004 (first entry)
XX XX

```

```

DE XX Rhesus theta defensin analogue peptide RTD-4.
XX KW Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial;
XX KW antiinflammatory; antibacterial; virucide; fungicide; food;
XX KW contact lens solution; eye wash solution; inflammatory response;
XX KW microbistatic inhibition; microbistatic growth inhibition; disinfectant;
XX KW food preservative; bacterial infection; viral infection;
XX KW fungal infection; haemolytic activity; cyclic.
XX OS Macaca mulatta.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1..18 "The peptide is cyclised by a covalent link
XX FT /note= between these two residues"
XX FT Disulfide-bond 3..16
XX FT Disulfide-bond 5..14
XX FT Disulfide-bond 7..12
XX PN US2004014669-A1.
XX XX
XX PD 22-JAN-2004.
XX PF 30-APR-2003; 2003US-00427715.
XX PR 30-APR-2002; 2002US-0377071P.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Selsted ME, Tran DQ;
XX DR WPI; 2004-167945/16.
XX PT Novel theta defensin analog useful for reducing or inhibiting growth or
XX PT survival of a microorganism in an environment such as food or food
XX PT product, contact lens solution, or eye wash solution, an inanimate
XX PT Object.
XX PS Claim 1; SEQ ID NO 29; 46pp; English.
XX CC The invention relates to a theta defensin analogue defined by formulae
XX CC detailed in the claims or appearing as ADO35239-ADO35257. The theta
XX CC defensin analogue is useful for reducing or inhibiting growth or survival
XX CC of a microorganism in an environment capable of sustaining the growth or
XX CC survival of the microorganism and is useful for reducing or inhibiting
XX CC growth or survival of a microorganism in an environment such as food or
XX CC food product, a solution (e.g., contact lens solution, or eye wash
XX CC solution), an inanimate object comprising surface, or a mammal. The
XX CC peptides are also useful for decreasing inflammatory response and for
XX CC microbistatic inhibition of survival of microorganism as well as
XX CC microbistatic inhibition of growth. Thus the peptides are useful as
XX CC therapeutic agents, disinfectants, food preservatives, or medicaments.
XX CC The peptides are also useful for treating a patient suffering from
XX CC bacterial, viral, fungal or other infection. The theta defensins have
XX CC high antimicrobial activity and low haemolytic activity. The present
XX CC sequence represents a Rhesus theta defensin analogue peptide.
XX SQ Sequence 18 AA;
XX Query Match 86.2%; Score 100; DB 8; Length 18;
XX Best Local Similarity 93.8%; Pred. No. 0.00038;
XX Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 1 GFCRCICTRGFCRCIC 16
XX Db ||||||
XX 1 GICRCICTRGFCRCIC 16
XX RESULT 18
XX AAE33805
XX ID AAE33805 standard; peptide; 18 AA.
XX XX

```

AAE33805;
 16-APR-2003 (first entry)
 12Y retrocyclin peptide analogue.
 Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
 human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
 antibiotic modelling; antimicrobial; cyclic.
 Unidentified.
 Key Location/Qualifiers
 Misc-difference 1 /note= "Linked to amino acid at position 18 to form a
 cyclic structure"
 Disulfide-bond 3. .16
 Disulfide-bond 5. .14
 Disulfide-bond 7. .12
 Misc-difference 18
 /note= "Linked to amino acid at position 1 to form a
 cyclic structure"
 WO200285401-A1.
 31-OCT-2002.
 18-APR-2002; 2002WO-US012353.
 18-APR-2001; 2001US-0284855P.
 (REGC) UNIV CALIFORNIA.
 Lehrer RI, Waring AJ, Cole AM, Hong TB;
 WPI; 2003-103387/09.
 New isolated retrocyclin peptide, useful for preventing retroviral
 infections in cells susceptible to bacterial or viral infections or
 treating patients having the infections, such as HIV, sexually
 transmitted diseases, vaginosis.
 Claim 9; Page 24; 72pp; English.
 The invention relates to novel retrocyclin peptides. Peptides and methods
 of the invention are useful for preventing retroviral infections in cells
 susceptible to bacterial or viral infections, or treating patients having
 infections such as HIV (human immunodeficiency virus), sexually
 transmitted diseases, bacterial vaginosis or ophthalmic infections. The
 retrocyclin-mediated killing is useful for modelling and screening novel
 antibiotics. The invention is also useful in gene therapy. The present
 sequence is human retrocyclin peptide analogue
 Sequence 18 AA;
 Query Match 80.2%; Score 93; DB 6; Length 18;
 Best Local Similarity 77.8%; Pred. No. 0.0021;
 Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GFCRCICTRGFCRCICTR 18
 |:|||||:|||||:
 Db 1 GYCRCICGRGICRCICGR 18
 RESULT 19
 AAE33806
 ID AAE33806 standard; peptide; 18 AA.
 AC AAE33806;
 16-APR-2003 (first entry)
 11ly retrocyclin peptide analogue.

XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
 KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
 KW antibiotic modelling; antimicrobial; cyclic.
 OS Unidentified.
 XX Key Location/Qualifiers
 FH Misc-difference 1 /note= "Linked to amino acid at position 18 to form a
 cyclic structure"
 FT Disulfide-bond 3. .16
 FT Disulfide-bond 5. .14
 FT Disulfide-bond 7. .12
 FT Misc-difference 18
 /note= "Linked to amino acid at position 1 to form a
 cyclic structure"
 WO200285401-A1.
 31-OCT-2002.
 18-APR-2002; 2002WO-US012353.
 18-APR-2001; 2001US-0284855P.
 (REGC) UNIV CALIFORNIA.
 Lehrer RI, Waring AJ, Cole AM, Hong TB;
 WPI; 2003-103387/09.
 New isolated retrocyclin peptide, useful for preventing retroviral
 infections in cells susceptible to bacterial or viral infections or
 treating patients having the infections, such as HIV, sexually
 transmitted diseases, vaginosis.
 Claim 9; Page 24; 72pp; English.
 The invention relates to novel retrocyclin peptides. Peptides and methods
 of the invention are useful for preventing retroviral infections in cells
 susceptible to bacterial or viral infections, or treating patients having
 infections such as HIV (human immunodeficiency virus), sexually
 transmitted diseases, bacterial vaginosis or ophthalmic infections. The
 retrocyclin-mediated killing is useful for modelling and screening novel
 antibiotics. The invention is also useful in gene therapy. The present
 sequence is human retrocyclin peptide analogue
 Sequence 18 AA;
 Query Match 80.2%; Score 93; DB 6; Length 18;
 Best Local Similarity 77.8%; Pred. No. 0.0021;
 Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GFCRCICTRGFCRCICTR 18
 |:|||||:|||||:
 Db 1 GYCRCICGRGICRCICGR 18
 RESULT 20
 ADN08180
 ID ADN08180 standard; peptide; 18 AA.
 AC ADN08180;
 XX
 DT 15-JUL-2004. (first entry)
 XX Human retrocyclin peptide, RC-104.
 DE retrocyclin; cyclic; primate; retroviral infection;
 KW circular minidefensin; antibacterial; virucide; vaccine; immunotherapy;
 KW microbial; viral; human.

OS Homo sapiens.
XX Synthetic.

FN WO2004033479-A2.

XX PD 22-APR-2004.

XX PF 06-MAY-2003; 2003WO-US014106.

XX PR 06-MAY-2002; 2002US-00141645.

XX PA (REGC) UNIV CALIFORNIA.

XX PA (LEHR/) LEHRER R I.

XX PA (WARI/) WARING A J.

XX PA (COLE/) COLE A M.

XX PA (HONG/) HONG T B.

XX PI Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX DR WPI; 2004-340883/31.

XX PS New isolated retrocyclin peptides and cyclic polypeptides, useful as

XX PT therapeutic and prophylactic agents for treating and preventing microbial

XX FT and viral infections.

XX PS Claim 9; SEQ ID NO 5; 82pp; English.

XX CC The invention relates to a novel isolated retrocyclin peptide. The

XX CC invention further provides: a cyclic polypeptide; an isolated nucleic

XX CC acid encoding a primate retrocyclin; a method for preventing retroviral

XX CC infection in a cell by administering an effective dose of a circular

XX CC minidensin or retrocyclin to the cell; a method for killing microbial

XX CC organisms by administering an effective dose of retrocyclin to the

XX CC microbial organisms; a method for administering retrocyclin as a

XX CC therapeutic agent to a patient with an established microbial or viral

XX CC infection; and a method for administering retrocyclin as a prophylactic

XX CC agent to prevent a microbial or viral infection in a patient at risk of

XX CC developing such infection. The retrocyclin peptide has antibacterial and

XX CC virucide activities. The retrocyclin peptide can be used to treat a viral

XX CC infection such as HIV-1. The retrocyclin peptide can be used as a vaccine

XX CC and can be used in immunotherapy. The peptide and polypeptides are useful

XX CC as therapeutic and prophylactic agents for treating and preventing

XX CC microbial and viral infections. This sequence represents a retrocyclin

XX CC peptide of the invention.

XX SQ Sequence 18 AA;

Query Match 80.2%; Score 93; DB 8; Length 18;

Best Local Similarity 77.8%; Pred. No. 0.0021;

Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18

Db 1 GYCRICGRCICRCICGR 18

RESULT 21

ADN08181

ID ADN08181 standard; peptide; 18 AA.

XX AC ADN08181;

XX DT 15-JUL-2004 (first entry)

XX DE Human retrocyclin peptide, RC-105.

XX KW retrocyclin; cyclic; primate; retroviral infection;

XX KW circular minidensin; antibacterial; virucide; vaccine; immunotherapy;

XX KW microbial; viral; human.

XX OS Homo sapiens.

XX OS Synthetic.

PN WO2004033479-A2.

XX PD 22-APR-2004.

XX PF 06-MAY-2003; 2003WO-US014106.

XX PR 06-MAY-2002; 2002US-00141645.

XX PA (REGC) UNIV CALIFORNIA.

XX PA (LEHR/) LEHRER R I.

XX PA (WARI/) WARING A J.

XX PA (COLE/) COLE A M.

XX PA (HONG/) HONG T B.

XX PI Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX DR WPI; 2004-340883/31.

XX PT New isolated retrocyclin peptides and cyclic polypeptides, useful as

XX FT therapeutic and prophylactic agents for treating and preventing microbial

XX PS Claim 9; SEQ ID NO 6; 82pp; English.

XX CC The invention relates to a novel isolated retrocyclin peptide. The

XX CC invention further provides: a cyclic polypeptide; an isolated nucleic

XX CC acid encoding a primate retrocyclin; a method for preventing retroviral

XX CC infection in a cell by administering an effective dose of a circular

XX CC minidensin or retrocyclin to the cell; a method for killing microbial

XX CC organisms by administering an effective dose of retrocyclin to the

XX CC microbial organisms; a method for administering retrocyclin as a

XX CC therapeutic agent to a patient with an established microbial or viral

XX CC infection; and a method for administering retrocyclin as a prophylactic

XX CC agent to prevent a microbial or viral infection in a patient at risk of

XX CC developing such infection. The retrocyclin peptide has antibacterial and

XX CC virucide activities. The retrocyclin peptide can be used to treat a viral

XX CC infection such as HIV-1. The retrocyclin peptide can be used as a vaccine

XX CC and can be used in immunotherapy. The peptide and polypeptides are useful

XX CC as therapeutic and prophylactic agents for treating and preventing

XX CC microbial and viral infections. This sequence represents a retrocyclin

XX CC peptide of the invention.

XX SQ Sequence 18 AA;

Query Match 80.2%; Score 93; DB 8; Length 18;

Best Local Similarity 77.8%; Pred. No. 0.0021;

Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18

Db 1 GICRCICGRCICRCICGR 18

RESULT 22

AAB35037

ID AAB35037 standard; peptide; 18 AA.

XX AC AAB35037;

XX DT 27-MAR-2001 (first entry)

XX DE Rhesus macaque theta defensin peptide SEQ ID NO: 8.

XX KW Theta defensin; antimicrobial; cyclic; bacterium; fungus; protozoan;

XX KW virus; helminth; disinfectant; food preservative; analogue.

XX OS Rhesus macaque.

XX OS WO200068265-A1.

XX PF 16-NOV-2000.

XX PF 10-MAY-2000; 2000WO-US012842.

XX 10-MAY-1999; 99US-00309487.
XX (REGC) UNIV CALIFORNIA.

XX Selsted ME, Tang Y, Yuan J, Ouellette AJ;

XX WPI; 2001-031853/04.

XX Novel theta defensin peptide with antimicrobial activity against
XX bacteria, yeast, fungi, protozoa and viruses.

XX Example 1; Fig 2; 110pp; English.

XX The present invention provides theta defensin peptides and analogues
XX which have antimicrobial activity. They can be used in the treatment of
XX bacterial, viral, fungal, protozoan and helminthic infections, in
XX disinfectants and as food preservatives

XX Sequence 18 AA;

Query Match 78.4%; Score 91; DB 4; Length 18;
Best Local Similarity 81.2%; Pred. No. 0.0033;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCIC 16

Db 3 GFCRCILCRGVCRIC 18

RESULT 23.

ID ABP53294
ID ABP53294 standard; peptide; 18 AA.

AC ABP53294;

DT 13-NOV-2002 (first entry)

XX Synthetic anti-viral human theta defensin peptide HTD-1 SEQ ID NO:27.

XX Anti-viral; viral infection; theta-defensin; lipid environment;
XX amphipathic alpha-helical structure; virucide; anti-HIV; immunisation;
XX viral growth inhibitor; viral proliferation inhibitor.

XX Homo sapiens.

XX Synthetic.

XX WO200260468-A2.

XX 08-AUG-2002.

XX 29-JAN-2002; 2002WO-US002435.

XX 30-JAN-2001; 2001US-0265270P.

XX 01-AUG-2001; 2001US-0309368P.

XX (IOWA) UNIV IOWA RES FOUND.

XX Maury W, Stapleton J, Stinski M, Roller R, Mcoray PB, Tack B;

XX WPI; 2002-674815/72.

XX New method of using a first anti-viral peptide comprising a Theta-
XX defensin peptide in an amphipathic Alpha-helical structure in a lipid
XX environment for reducing the infectivity of a virus.

XX Disclosure; Page 9; 65pp; English.

XX The present invention describes a method (M1) of using a first anti-viral
XX peptide (I) comprising a theta-defensin peptide in an amphipathic alpha-
XX helical structure in a lipid environment for reducing the infectivity of
XX a virus. (I) can have virucide and anti-HIV activities, and can be used
XX to reduce virus growth, infectivity burden, shed, and development of anti

CC -viral resistance. (I) can be used for inhibiting the growth and
CC proliferation of a virus and so can be used for; (a) protecting or
CC treating subject from a viral infection, preventing recurrent viral
CC infection in a subject harbouring a latent virus, controlling virus
CC spread within a virally-infected subject (VS), reducing viral burden in a
CC VS, reducing virus shed from a VS, reducing percentage of VS in a
CC population regardless of viral infection status, or inducing latency in a
CC VS; (b) reducing the infectivity of a virus; and (c) reducing the number
CC of contaminated tissue or fluid sample safe for use, or reducing the number
CC of infectious virus particles in a population of viruses. (M1) is useful
CC for reducing the infectivity of a virus in sheep, cattle, horses, swine,
CC cats, fowl and humans e.g. an enveloped virus infecting humans such as
CC human immunodeficiency virus (HIV). Preferably, the anti-viral peptide is
CC administered to a patient who is immunosuppressed or to a subject who is
CC not infected with the virus, where the first anti-viral peptide is
CC administered prior to or subsequent to the virus contacting the subject.
CC The anti-viral peptide is most preferably administered to a subject who
CC is chronically, latently or acutely infected with the virus. The present
CC sequence represents a human theta defensin anti-viral peptide, which is
CC given in the exemplification of the present invention

XX Sequence 18 AA;

Query Match 77.6%; Score 90; DB 5; Length 18;

Best Local Similarity 77.8%; Pred. No. 0.0043;

Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCIC 18

Db 1 GICRCICRGICRCIC 18

RESULT 24

AAE33801

ID AAE33801 standard; peptide; 18 AA.

XX AAE33801;

XX 16-APR-2003 (first entry)

XX Human retrocyclin peptide.

XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;

XX human immunodeficiency virus; bacterial vaginosis; opththalmic infection;
XX antibiotic modelling; antimicrobial; human; cyclic.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 1 /note= "Linked to amino acid at position 18 to form a
XX cyclic structure"

XX Disulfide-bond 3. .16

XX Disulfide-bond 5. .14

XX Disulfide-bond 7. .12

XX Misc-difference 18

XX /note= "Linked to amino acid at position 1 to form a
XX cyclic structure"

XX WO200285401-A1.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012353.

XX 18-APR-2001; 2001US-0284855P.

XX (REGC) UNIV CALIFORNIA.

XX Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2003-103387/09.

PT New isolated retrocyclin peptide, useful for preventing retroviral
PT infections in cells susceptible to bacterial or viral infections or
PT treating patients having the infections, such as HIV, sexually
PT transmitted diseases, vaginosis.
XX
PS Claim 9; Page 24; 72pp; English.
XX
CC The invention relates to novel retrocyclin peptides. Peptides and methods
CC of the invention are useful for preventing retroviral infections in cells
CC susceptible to bacterial or viral infections, or treating patients having
CC infections such as HIV (human immunodeficiency virus), sexually
CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
CC retrocyclin-mediated killing is useful for modelling and screening novel
CC antibiotics. The invention is also useful in gene therapy. The present
CC sequence is human retrocyclin peptide
XX
SQ Sequence 18 AA;
Query Match 77.6%; Score 90; DB 6; Length 18;
Best Local Similarity 77.8%; Pred. No. 0.0043;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 GFCRCICTRGFCRCICTR 18
Db 1 GICRCICGRCICRCICGR 18
RESULT 25
AAE33863
ID AAE33863 standard; peptide; 18 AA.
XX
AC AAE33863;
XX
DT 16-APR-2003 (first entry)
XX
DE Enantio-retrocyclin peptide analogue.
XX
KW Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
KW antibiotic modelling; antimicrobial; cyclic.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .18
FT /note= "D-form residues"
FT Misc-difference 1
FT /note= "Linked to amino acid at position 18 to form a
FT cyclic structure"
FT Disulfide-bond 3. .16
FT Disulfide-bond 5. .14
FT Disulfide-bond 7. .12
FT Misc-difference 18
FT /note= "Linked to amino acid at position 1 to form a
FT cyclic structure"
XX
PN WO200285401-A1.
XX
PD 31-OCT-2002.
XX
PF 18-APR-2002; 2002WO-US012353.
XX
PR 18-APR-2001; 2001US-0284855P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Lehrer RI, Waring AJ, Cole AM, Hong TB;
XX
DR WPI; 2003-103387/09.
XX
CC New isolated retrocyclin peptide, useful for preventing retroviral
PT infections in cells susceptible to bacterial or viral infections or
PT treating patients having the infections, such as HIV, sexually

PT transmitted diseases, vaginosis.
XX
PS Disclosure; Page 24; 72pp; English.
XX
CC The invention relates to novel retrocyclin peptides. Peptides and methods
CC of the invention are useful for preventing retroviral infections in cells
CC susceptible to bacterial or viral infections, or treating patients having
CC infections such as HIV (human immunodeficiency virus), sexually
CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
CC retrocyclin-mediated killing is useful for modelling and screening novel
CC antibiotics. The invention is also useful in gene therapy. The present
CC sequence is human retrocyclin peptide analogue
XX
SQ Sequence 18 AA;
Query Match 77.6%; Score 90; DB 6; Length 18;
Best Local Similarity 77.8%; Pred. No. 0.0043;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 GFCRCICTRGFCRCICTR 18
Db 1 GICRCICGRCICRCICGR 18
RESULT 26
ADN08176
ID ADN08176 standard; peptide; 18 AA.
XX
AC ADN08176;
XX
DT 15-JUL-2004 (first entry)
XX
DE Human retrocyclin peptide, RC-100.
XX
KW retrocyclin; cyclic; primate; retroviral infection;
KW circular minidensin; antibacterial; virucide; vaccine; immunotherapy;
KW microbial; viral; human.
XX
OS Homo sapiens.
XX
PN WO2004033479-A2.
XX
PD 22-APR-2004.
XX
PF 06-MAY-2003; 2003WO-US014106.
XX
PR 06-MAY-2002; 2002US-00141645.
XX
PA (REGC) UNIV CALIFORNIA.
PA (LEHR/) LEHRER R I.
PA (WARI/) WARING A J.
PA (COLE/) COLE A W.
PA (HONG/) HONG T B.
XX
PI Lehrer RI, Waring AJ, Cole AM, Hong TB;
XX
DR WPI; 2004-340883/31.
DR N-PSDB; ADN08193.
XX
PT New isolated retrocyclin peptides and cyclic polypeptides, useful as
PT therapeutic and prophylactic agents for treating and preventing microbial
PT and viral infections.
XX
PS Claim 9; SEQ ID NO 1; 82pp; English.
XX
CC The invention relates to a novel isolated retrocyclin peptide. The
CC invention further provides: a cyclic polypeptide; an isolated nucleic
CC acid encoding a primate retrocyclin; a method for preventing retroviral
CC infection in a cell by administering an effective dose of a circular
CC minidensin or retrocyclin to the cell; a method for killing microbial
CC organisms by administering an effective dose of retrocyclin to the
CC microbial organisms; a method for administering retrocyclin as a
CC therapeutic agent to a patient with an established microbial or viral

CC infection; and a method for administering retrocyclin as a prophylactic
 CC agent to prevent a microbial or viral infection in a patient at risk of
 CC developing such infection. The retrocyclin peptide has antibacterial and
 CC virucide activities. The retrocyclin peptide can be used to treat a viral
 CC infection such as HIV-1. The retrocyclin peptide can be used as a vaccine
 CC and can be used in immunotherapy. The peptide and polypeptides are useful
 CC as therapeutic and prophylactic agents for treating and preventing
 CC microbial and viral infections. This sequence represents a retrocyclin
 CC peptide of the invention.

XX Sequence 18 AA;
 SQ Query Match 77.6%; Score 90; DB 8; Length 18;
 Best Local Similarity 77.8%; Pred. No. 0.0043;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
 DQ 1 GFCRCICTRGFCRCICTR 18

RESULT 27
 ADO35249
 ID ADO35249 standard; peptide; 18 AA.

XX AC ADO35249;
 XX DT 15-JUL-2004 (first entry)
 XX DE Rhesus theta defensin analogue peptide RTD-1-25.

XX Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial;
 KW antiinflammatory; antibacterial; virucide; fungicide; food;
 KW contact lens solution; eye wash solution; inflammatory response;
 KW microbicidal inhibition; microbistatic growth inhibition; disinfectant;
 KW food preservative; bacterial infection; viral infection;
 KW fungal infection; haemolytic activity; cyclic.

XX Macaca mulatta.
 OS Synthetic.

PH Key Location/Qualifiers
 FT Modified-site 1..18
 FT /note= "The peptide is cyclised by a covalent link
 FT between these two residues"

FT Disulfide-bond 3..16
 FT Disulfide-bond 5..14
 FT Disulfide-bond 7..12

XX US2004014669-A1.

XX 22-JAN-2004.

XX 30-APR-2003; 2003US-00427715.

XX 30-APR-2002; 2002US-0377071P.

XX (REGC) UNIV CALIFORNIA.

XX Selsted ME, Tran DQ;

XX WPI; 2004-167945/16.

XX Novel theta defensin analog useful for reducing or inhibiting growth or
 PT survival of a microorganism in an environment such as food or food
 PT product, contact lens solution, or eye wash solution, an inanimate
 PT object.

XX Claim 1; SEQ ID NO 23; 46pp; English.

XX The invention relates to a theta defensin analogue defined by formulae
 CC detailed in the claims or appearing as ADO35239-ADO35257. The theta
 CC defensin analogue is useful for reducing or inhibiting growth or survival

CC of a microorganism in an environment capable of sustaining the growth or
 CC survival of the microorganism and is useful for reducing or inhibiting
 CC growth or survival of a microorganism in an environment such as food or
 CC food product, a solution (e.g., contact lens solution, or eye wash
 CC solution), an inanimate object comprising surface, or a mammal. The
 CC peptides are also useful for decreasing inflammatory response and for
 CC microbicidal inhibition of survival of microorganism as well as
 CC microbistatic inhibition of growth. Thus the peptides are useful as
 CC therapeutic agents, disinfectants, food preservatives, or medicaments.
 CC The peptides are also useful for treating a patient suffering from
 CC bacterial, viral, fungal or other infection. The theta defensins have
 CC high antimicrobial activity and low haemolytic activity. The present
 CC sequence represents a Rhesus theta defensin analogue peptide.

XX Sequence 18 AA;

Query Match 76.7%; Score 89; DB 8; Length 18;
 Best Local Similarity 77.8%; Pred. No. 0.0054;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
 DQ 1 GFCRCICTRGFCRCICTR 18

RESULT 28
 ABP53298
 ID ABP53298 standard; peptide; 18 AA.

XX AC ABP53298;

XX DT 13-NOV-2002 (first entry)

XX Anti-viral chimeric theta defensin peptide H/RTD-3 SEQ ID NO:31.

DE Anti-viral; viral infection; theta-defensin; lipid environment;
 XX amphipathic alpha-helical structure; virucide; anti-HIV; immunisation;
 KW viral growth inhibitor; viral proliferation inhibitor.

OS Homo sapiens.
 OS Macaca mulatta.
 OS Synthetic.

XX WO2002060468-A2.

XX 08-AUG-2002.

XX 29-JAN-2002; 2002WO-US002435.

XX 30-JAN-2001; 2001US-0265270P.

XX 01-AUG-2001; 2001US-0309368P.

XX (IOWA) UNIV IOWA RES FOUND.

XX Maury W, Stapleton J, Stinski M, Roller R, McCray PB, Tack B;
 XX WPI; 2002-674815/72.

XX New method of using a first anti-viral peptide comprising a Theta-
 PT defensin peptide in an amphipathic Alpha-helical structure in a lipid
 PT environment for reducing the infectivity of a virus.

XX Disclosure; Page 10; 65pp; English.

XX The present invention describes a method (M1) of using a first anti-viral
 CC peptide (I) comprising a theta-defensin peptide in an amphipathic alpha-
 CC helical structure in a lipid environment for reducing the infectivity of
 CC a virus. (I) can have virucide and anti-HIV activities, and can be used
 CC to reduce virus growth, infectivity burden, shed, and development of anti-
 CC -viral resistance. (I) can be used for inhibiting the growth and
 CC proliferation of a virus and so can be used for; (a) protecting or
 CC treating subject from a viral infection, preventing recurrent viral
 CC infection in a subject harbouring a latent virus, controlling virus

CC The invention relates to novel retrocyclin peptides. Peptides and methods
 CC of the invention are useful for preventing retroviral infections in cells
 CC susceptible to bacterial or viral infections, or treating patients having
 CC infections such as HIV (human immunodeficiency virus), sexually
 CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
 CC retrocyclin-mediated killing is useful for modelling and screening novel
 CC antibiotics. The invention is also useful in gene therapy. The present
 CC sequence is human retrocyclin peptide analogue

XX Sequence 18 AA;

Query Match 75.0%; Score 87; DB 6; Length 18;
 Best Local Similarity 72.2%; Pred. No. 0.0088;
 Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
 Db 1 GICRCICGICRCICGR 18

RESULT 31

ID AAE33802 standard; peptide; 18 AA.

XX AAE33802;

DT 16-APR-2003 (first entry)

DE R9K retrocyclin peptide analogue.

KW Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
 KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
 KW antibiotic modelling; antimicrobial; cyclic.

XX Unidentified.

Key Location/Qualifiers

FT Misc-difference 1 /note= "linked to amino acid at position 18 to form a
 FT cyclic structure"
 FT Disulfide-bond 3. .16
 FT Disulfide-bond 5. .14
 FT Disulfide-bond 7. .12
 FT Misc-difference 18 /note= "linked to amino acid at position 1 to form a
 FT cyclic structure"

XX WO200285401-A1.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012353.

XX 18-APR-2001; 2001US-0284855P.

XX (REGC) UNIV CALIFORNIA.

XX Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2003-103387/09.

XX New isolated retrocyclin peptide, useful for preventing retroviral
 PT infections in cells susceptible to bacterial or viral infections or
 PT treating patients having the infections, such as HIV, sexually
 PT transmitted diseases, vaginosis.

XX Claim 9; Page 24; 72pp; English.

XX The invention relates to novel retrocyclin peptides. Peptides and methods
 CC of the invention are useful for preventing retroviral infections in cells
 CC susceptible to bacterial or viral infections, or treating patients having
 CC infections such as HIV (human immunodeficiency virus), sexually
 CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The

CC retrocyclin-mediated killing is useful for modelling and screening novel
 CC antibiotics. The invention is also useful in gene therapy. The present
 CC sequence is human retrocyclin peptide analogue

XX Sequence 18 AA;

Query Match 75.0%; Score 87; DB 6; Length 18;
 Best Local Similarity 72.2%; Pred. No. 0.0088;
 Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
 Db 1 GICRCICGICRCICGR 18

RESULT 32

ADN08177
 ID ADN08177 standard; peptide; 18 AA.

XX ADN08177;

DT 15-JUL-2004 (first entry)

DE Human retrocyclin peptide, RC-101.

KW retrocyclin; cyclic; primate; retroviral infection;
 KW circular minidefensin; antibacterial; virucide; vaccine; immunotherapy;
 KW microbial; viral; human.

XX Homo sapiens.

OS Synthetic.

XX WO2004033479-A2.

XX 22-APR-2004.

XX 06-MAY-2003; 2003WO-US014106.

XX 06-MAY-2002; 2002US-00141645.

XX (REGC) UNIV CALIFORNIA.

XX (LEHR/) LEHRER R I.

XX (WARI/) WARING A J.

XX (COLE/) COLE A M.

XX (HONG/) HONG T B.

XX Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2004-340883/31.

XX New isolated retrocyclin peptides and cyclic polypeptides, useful as
 PT therapeutic and prophylactic agents for treating and preventing microbial
 PT and viral infections.

XX Claim 9; SEQ ID NO 2; 82pp; English.

XX The invention relates to a novel isolated retrocyclin peptide. The
 CC invention further provides: a cyclic polypeptide; an isolated nucleic
 CC acid encoding a primate retrocyclin; a method for preventing retroviral
 CC infection in a cell by administering an effective dose of a circular
 CC minidefensin or retrocyclin to the cell; a method for killing microbial
 CC organisms by administering an effective dose of retrocyclin to the
 CC microbial organisms; a method for administering retrocyclin as a
 CC therapeutic agent to a patient with an established microbial or viral
 CC infection; and a method for administering retrocyclin as a prophylactic
 CC agent to prevent a microbial or viral infection in a patient at risk of
 CC developing such infection. The retrocyclin peptide has antibacterial and
 CC virucide activities. The retrocyclin peptide can be used to treat a viral
 CC infection such as HIV-1. The retrocyclin peptide can be used as a vaccine
 CC and can be used in immunotherapy. The peptide and polypeptides are useful
 CC as therapeutic and prophylactic agents for treating and preventing
 CC microbial and viral infections. This sequence represents a retrocyclin
 CC peptide of the invention.

XX SQ Sequence 18 AA;
Query Match 75.0%; Score 87; DB 8; Length 18;
Best Local Similarity 72.2%; Pred. No. 0.0088;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 GFCRCICTRGFCRCICTR 18
DB 1 GICRCICGRCICRCICGR 18
RESULT 33
ID AAB35047
AC AAB35047; peptide; 18 AA.
XX DT 27-MAR-2001 (first entry)
XX DE Theta defensin SEQ ID NO: 31.
XX KW Theta defensin; antimicrobial; cyclic; bacterium; fungus; protozoan;
XX KW virus; helminth; disinfectant; food preservative; analogue.
XX OS Unidentified.
XX PN WO200068265-A1.
XX PD 16-NOV-2000.
XX PF 10-MAY-2000; 2000WO-US012842.
XX PR 10-MAY-1999; 99US-00309487.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Selsted ME, Tang Y, Yuan J, Ouellette AJ;
XX WPI; 2001-031853/04.
XX PT Novel theta defensin peptide with antimicrobial activity against
XX PT bacteria, yeast, fungi, protozoa and viruses.
XX PS Claim 19, Fig 16; 11Opp; English.
XX CC The present invention provides theta defensin peptides and analogues
XX CC which have antimicrobial activity. They can be used in the treatment of
XX CC bacterial, viral, fungal, protozoan and helminthic infections, in
XX CC disinfectants and as food preservatives
XX SQ Sequence 18 AA;
Query Match 74.1%; Score 86; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.011;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 GFCRCICTRGFCRCICTR 18
DB 1 GVCRCICRGVCRCICRR 18
RESULT 34
ID ABP53296
AC ABP53296; peptide; 18 AA.
XX DT 13-NOV-2002 (first entry)
XX DE Anti-viral theta defensin peptide RTD-3 SEQ ID NO:29.
XX KW Anti-viral; viral infection; theta-defensin; lipid environment;

KW amphipathic alpha-helical structure; virucide; anti-HIV; immunisation;
KW viral growth inhibitor; viral proliferation inhibitor.
OS Macaca mulatta.
OS Synthetic.
XX PN WO200260468-A2.
XX PD 08-AUG-2002.
XX PF 29-JAN-2002; 2002WO-US002435.
XX PR 30-JAN-2001; 2001US-0265270P.
XX PR 01-AUG-2001; 2001US-0309368P.
XX PA (IOWA) UNIV IOWA RES FOUND.
XX PI Maury W, Stapleton J, Stinski M, Roller R, McCreay PB, Tack B;
XX WPI; 2002-674815/72.
XX DR
XX PT New method of using a first anti-viral peptide comprising a Theta-
XX PT defensin peptide in an amphipathic Alpha-helical structure in a lipid
XX PT environment for reducing the infectivity of a virus.
XX PS Disclosure; Page 10; 65pp; English.
XX CC The present invention describes a method (M1) of using a first anti-viral
XX CC peptide (I) comprising a theta-defensin peptide in an amphipathic alpha-
XX CC helical structure in a lipid environment for reducing the infectivity of
XX CC a virus. (I) can have virucide and anti-HIV activities, and can be used
XX CC to reduce virus growth, infectivity burden, shed, and development of anti-
XX CC -viral resistance. (I) can be used for inhibiting the growth and
XX CC proliferation of a virus and so can be used for: (a) protecting or
XX CC treating subject from a viral infection, preventing recurrent viral
XX CC infection in a subject harbouring a latent virus, controlling virus
XX CC spread within a virally-infected subject (VS), reducing viral burden in a
XX CC VS, reducing virus shed from a VS, reducing percentage of VS in a
XX CC population regardless of viral infection status, or inducing latency in a
XX CC VS; (b) reducing the infectivity of a virus; and (c) rendering virus-
XX CC contaminated tissue or fluid sample safe for use, or reducing the number
XX CC of infectious virus particles in a population of viruses. (M1) is useful
XX CC for reducing the infectivity of a virus in sheep, cattle, horses, swine,
XX CC cats, fowl and humans e.g. an enveloped virus infecting humans such as
XX CC human immunodeficiency virus (HIV). Preferably, the anti-viral peptide is
XX CC administered to a patient who is immunosuppressed or to a subject who is
XX CC not infected with the virus, where the first anti-viral peptide is
XX CC administered prior to or subsequent to the virus contacting the subject.
XX CC The anti-viral peptide is most preferably administered to a subject who
XX CC is chronically, latently or acutely infected with the virus. The present
XX CC sequence represents a rhesus monkey theta defensin anti-viral peptide,
XX CC which is given in the exemplification of the present invention
XX SQ Sequence 18 AA;
Query Match 74.1%; Score 86; DB 5; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.011;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 GFCRCICTRGFCRCICTR 18
DB 1 GVCRCICRGVCRCICRR 18
RESULT 35
ID ADO35231
AC ADO35231; standard; peptide; 18 AA.
XX DT 15-JUL-2004 (first entry)
XX DE Rhesus theta defensin peptide, RTD-3.

XX KW Monkey; Rhesus theta defensin; RTD-3; antimicrobial peptide; cyclic;
 XX KW antimicrobial; antiinflammatory; antibacterial; virucide; fungicide;
 XX KW food; contact lens solution; eye wash solution; inflammatory response;
 XX KW microbicidal inhibition; microbistatic growth inhibition; disinfectant;
 XX KW food preservative; bacterial infection; viral infection;
 XX KW fungal infection; haemolytic activity.
 OS Macaca mulatta.
 XX Key Location/Qualifiers
 XX Modified-site 1. .18
 FT /note= "The peptide is cyclised by a covalent link
 FT between these two residues"
 FT Disulfide-bond 3. .16
 FT Disulfide-bond 5. .14
 FT Disulfide-bond 7. .12
 XX US2004014669-A1.
 XX 22-JAN-2004.
 XX 30-APR-2003; 2003US-00427715.
 XX 30-APR-2002; 2002US-0377071P.
 XX (REGC) UNIV CALIFORNIA.
 XX Selsted ME, Tran DQ;
 XX WPI; 2004-167945/16.
 XX Novel theta defensin analog useful for reducing or inhibiting growth or
 XX survival of a microorganism in an environment such as food or food
 XX product, contact lens solution, or eye wash solution, an inanimate
 XX object.
 PS Example 1; SEQ ID NO 3; 46pp; English.
 XX The invention relates to a theta defensin analogue defined by formulae
 XX detailed in the claims or appearing as ADO35239-ADO35257. The theta
 XX defensin analogue is useful for reducing or inhibiting growth or survival
 XX of a microorganism in an environment capable of sustaining the growth or
 XX survival of the microorganism and is useful for reducing or inhibiting
 XX growth or survival of a microorganism in an environment such as food or
 XX food product, a solution (e.g., contact lens solution, or eye wash
 XX solution), an inanimate object comprising surface, or a mammal. The
 XX peptides are also useful for decreasing inflammatory response and for
 XX microbicidal inhibition of survival of microorganism as well as
 XX microbistatic inhibition of growth. Thus the peptides are useful as
 XX therapeutic agents, disinfectants, food preservatives, or medicaments.
 XX The peptides are also useful for treating a patient suffering from
 XX bacterial, viral, fungal or other infection. The theta defensins have
 XX high antimicrobial activity and low haemolytic activity. The present
 XX sequence represents the rhesus monkey wild-type theta defensin RTD-3.
 SQ Sequence 18 AA;
 Query Match 74.1%; Score 86; DB 8; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.011;
 Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GFCRCICTRGFCRCICTR 18
 | | | | | | | | | | | | | | | | | |
 Db 1 GVCRCICRRGVCRCLCRR 18
 RESULT 36
 ADO35242
 ID ADO35242 standard; peptide; 18 AA.
 XX AC ADO35242;
 XX

DT 15-JUL-2004 (first entry)
 XX Rhesus theta defensin analogue peptide arTD-3-NH.
 DE Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial;
 XX KW antiinflammatory; antibacterial; virucide; fungicide; food;
 KW contact lens solution; eye wash solution; inflammatory response;
 KW microbicidal inhibition; microbistatic growth inhibition; disinfectant;
 KW food preservative; bacterial infection; viral infection;
 KW fungal infection; haemolytic activity.
 XX Macaca mulatta.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT Disulfide-bond 3. .16
 FT Disulfide-bond 5. .14
 FT Disulfide-bond 7. .12
 FT Modified-site 18
 FT /note= "Amidated"
 XX US2004014669-A1.
 XX 22-JAN-2004.
 XX 30-APR-2003; 2003US-00427715.
 XX 30-APR-2002; 2002US-0377071P.
 XX (REGC) UNIV CALIFORNIA.
 XX Selsted ME, Tran DQ;
 XX WPI; 2004-167945/16.
 XX Novel theta defensin analog useful for reducing or inhibiting growth or
 XX survival of a microorganism in an environment such as food or food
 XX product, contact lens solution, or eye wash solution, an inanimate
 XX object.
 PS Claim 1; SEQ ID NO 16; 46pp; English.
 XX The invention relates to a theta defensin analogue defined by formulae
 XX detailed in the claims or appearing as ADO35239-ADO35257. The theta
 XX defensin analogue is useful for reducing or inhibiting growth or survival
 XX of a microorganism in an environment capable of sustaining the growth or
 XX survival of the microorganism and is useful for reducing or inhibiting
 XX growth or survival of a microorganism in an environment such as food or
 XX food product, a solution (e.g., contact lens solution, or eye wash
 XX solution), an inanimate object comprising surface, or a mammal. The
 XX peptides are also useful for decreasing inflammatory response and for
 XX microbicidal inhibition of survival of microorganism as well as
 XX microbistatic inhibition of growth. Thus the peptides are useful as
 XX therapeutic agents, disinfectants, food preservatives, or medicaments.
 XX The peptides are also useful for treating a patient suffering from
 XX bacterial, viral, fungal or other infection. The theta defensins have
 XX high antimicrobial activity and low haemolytic activity. The present
 XX sequence represents a Rhesus theta defensin analogue peptide.
 SQ Sequence 18 AA;
 Query Match 74.1%; Score 86; DB 8; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.011;
 Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GFCRCICTRGFCRCICTR 18
 | | | | | | | | | | | | | | | | | |
 Db 1 GVCRCICRRGVCRCLCRR 18
 RESULT 37
 ADO35241
 ID ADO35241 standard; peptide; 18 AA.

XX ADO35241;
XX AC
XX 15-JUL-2004 (first entry)
XX DT
XX DE
XX Rhesus theta defensin analogue peptide ARTD-3-OH.
XX KW
XX Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial;
XX antinflammatory; antibacterial; virucide; fungicide; food;
XX contact lens solution; eye wash solution; inflammatory response;
XX microbicidal inhibition; microbistatic growth inhibition; disinfectant;
XX food preservative; bacterial infection; viral infection;
XX fungal infection; haemolytic activity.
XX OS
XX Macaca mulatta.
XX Synthetic.
XX XX
XX Key Location/Qualifiers
XX Disulfide-bond 3. .16
XX Disulfide-bond 5. .14
XX Disulfide-bond 7. .12
XX Modified-site 18
XX /note= "Hydroxylated"
XX US2004014669-A1.
XX PN
XX 22-JAN-2004.
XX PD
XX 30-APR-2003; 2003US-00427715.
XX PF
XX 30-APR-2002; 2002US-0377071P.
XX PR
XX (REGC) UNIV CALIFORNIA.
XX PA
XX Selsted ME, Tran DQ;
XX PI
XX WPI; 2004-167945/16.
XX DR
XX Novel theta defensin analog useful for reducing or inhibiting growth or
XX survival of a microorganism in an environment such as food or food
XX product, contact lens solution, or eye wash solution, an inanimate
XX object.
XX PT
XX Claim 1; SEQ ID NO 15; 46pp; English.
XX PS
XX The invention relates to a theta defensin analogue defined by formulae
XX detailed in the claims or appearing as ADO35239-ADO35257. The theta
XX defensin analogue is useful for reducing or inhibiting growth or survival
XX of a microorganism in an environment capable of sustaining the growth or
XX survival of the microorganism and is useful for reducing or inhibiting
XX growth or survival of a microorganism in an environment such as food or
XX food product, a solution (e.g., contact lens solution, or eye wash
XX solution), an inanimate object comprising surface, or a mammal. The
XX peptides are also useful for decreasing inflammatory response and for
XX microbicidal inhibition of survival of microorganism as well as
XX microbistatic inhibition of growth. Thus the peptides are useful as
XX therapeutic agents, disinfectants, food preservatives, or medicaments.
XX The peptides are also useful for treating a patient suffering from
XX bacterial, viral, fungal or other infection. The theta defensins have
XX high antimicrobial activity and low haemolytic activity. The present
XX sequence represents a Rhesus theta defensin analogue peptide.
XX XX
XX Sequence 18 AA;

Query Match 74.1%; Score 86; DB 8; Length 18;
Best Local Similarity 56.7%; Pred. No. 0.011;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GFCRCICTGFCRCICTR 18
Db 1 GVCRCICLRGVNCRICRR 18

RESULT 38
AAE33804
ID AAE33804 standard; peptide; 18 AA.
XX AC
XX AAE33804;
XX DT
XX 16-APR-2003 (first entry)
XX DE
XX I15Y retrocyclin peptide analogue.
XX KW
XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
XX human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
XX antibiotic modelling; antimicrobial; cyclic.
XX OS
XX Unidentified.
XX PH
XX Key Location/Qualifiers
XX Misc-difference 1 /note= "Linked to amino acid at position 18 to form a
XX cyclic structure"
XX Disulfide-bond 3. .16
XX Disulfide-bond 5. .14
XX Disulfide-bond 7. .12
XX Misc-difference 18
XX /note= "Linked to amino acid at position 1 to form a
XX cyclic structure"
XX WO200285401-A1.
XX PN
XX 31-OCT-2002.
XX PD
XX 18-APR-2002; 2002WO-US012353.
XX PF
XX 18-APR-2001; 2001US-0284855P.
XX PR
XX (REGC) UNIV CALIFORNIA.
XX PA
XX Lehrer RI, Waring AJ, Cole AM, Hong TB;
XX WPI; 2003-103387/09.
XX DR
XX New isolated retrocyclin peptide, useful for preventing retroviral
XX infections in cells susceptible to bacterial or viral infections or
XX treating patients having the infections, such as HIV, sexually
XX transmitted diseases, vaginosis.
XX PT
XX Claim 9; Page 24; 72pp; English.
XX PS
XX The invention relates to novel retrocyclin peptides. Peptides and methods
XX of the invention are useful for preventing retroviral infections in cells
XX susceptible to bacterial or viral infections, or treating patients having
XX infections such as HIV (human immunodeficiency virus), sexually
XX transmitted diseases, bacterial vaginosis or ophthalmic infections. The
XX retrocyclin-mediated killing is useful for modelling and screening novel
XX antibiotics. The invention is also useful in gene therapy. The present
XX sequence is human retrocyclin peptide analogue
XX XX
XX Sequence 18 AA;

Query Match 73.3%; Score 85; DB 6; Length 18;
Best Local Similarity 72.2%; Pred. No. 0.014;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GFCRCICTGFCRCICTR 18
Db 1 GICRCICRGICRCYGR 18

RESULT 39
AAE33803
ID AAE33803 standard; peptide; 18 AA.
XX AC
XX AAE33803;

XX DT 16-APR-2003 (first entry)
 XX DE 16V retrocyclin peptide analogue.
 XX KW Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
 KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
 KW antibiotic modelling; antimicrobial; cyclic.
 XX OS Unidentified.

XX Key Location/Qualifiers
 XX Key Modified-site 1. .18
 FT Misc-difference 1 /note= "linked to amino acid at position 18 to form a
 FT cyclic structure"
 FT Disulfide-bond 3. .16
 FT Disulfide-bond 5. .14
 FT Disulfide-bond 7. .12
 FT Misc-difference 18
 FT /note= "linked to amino acid at position 1 to form a
 FT cyclic structure"
 XX

XX WO200285401-A1.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012353.

XX 18-APR-2001; 2001US-0284855P.

XX (REGC) UNIV CALIFORNIA.

XX Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2003-103387/09.

XX New isolated retrocyclin peptide, useful for preventing retroviral
 PT infections in cells susceptible to bacterial or viral infections or
 PT treating patients having the infections, such as HIV, sexually
 PT transmitted diseases, vaginosis.

XX Claim 9; Page 24; 72pp; English.

XX The invention relates to novel retrocyclin peptides. Peptides and methods
 CC of the invention are useful for preventing retroviral infections in cells
 CC susceptible to bacterial or viral infections, or treating patients having
 CC infections such as HIV (human immunodeficiency virus), sexually
 CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
 CC retrocyclin-mediated killing is useful for modelling and screening novel
 CC antibiotics. The invention is also useful in gene therapy. The present
 CC sequence is human retrocyclin peptide analogue

XX Sequence 18 AA;

Query Match 73.3%; Score 85; DB 6; Length 18;

Best Local Similarity 72.2%; Pred. No. 0.014; Mismatches 5; Indels 0; Gaps 0;

Matches 13; Conservative 0;

QY 1 GFCRCICTRGFCRCICTR 18

Db 1 GICRCYCGRGICRCICGR 18

RESULT 40

ADO35256

ID ADO35256 standard; peptide; 18 AA.

XX ADO35256;

XX 15-JUL-2004 (first entry)

XX Rhesus theta defensin analogue peptide RTD-5.

XX

KW Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial;
 KW antiinflammatory; antibacterial; virucide; fungicide; food;
 KW contact lens solution; eye wash solution; inflammatory response;
 KW microbicidal inhibition; microbiostatic growth inhibition; disinfectant;
 KW food preservative; bacterial infection; viral infection;
 XX fungal infection; haemolytic activity; cyclic.
 OS Macaca mulatta.
 OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1. .18
 FT /note= "The peptide is cyclised by a covalent link
 FT between these two residues"

FT Disulfide-bond 3. .16

FT Disulfide-bond 5. .14

FT Disulfide-bond 7. .12

XX US2004014669-A1.

XX 22-JAN-2004.

XX 30-APR-2003; 2003US-00427715.

XX 30-APR-2002; 2002US-0377071P.

XX (REGC) UNIV CALIFORNIA.

XX Selsted ME, Tran DQ;

XX WPI; 2004-167945/16.

XX Novel theta defensin analog useful for reducing or inhibiting growth or
 PT survival of a microorganism in an environment such as food or food
 PT product, contact lens solution, or eye wash solution, an inanimate
 PT object.

XX Claim 1; SEQ ID NO 30; 46pp; English.

XX The invention relates to a theta defensin analogue defined by formulae
 CC detailed in the claims or appearing as ADO35239-ADO35257. The theta
 CC defensin analogue is useful for reducing or inhibiting growth or survival
 CC of a microorganism in an environment capable of sustaining the growth or
 CC survival of the microorganism and is useful for reducing or inhibiting
 CC growth or survival of a microorganism in an environment such as food or
 CC food product, a solution (e.g., contact lens solution, or eye wash
 CC solution), an inanimate object comprising surface, or a mammal. The
 CC peptides are also useful for decreasing inflammatory response and for
 CC microbicidal inhibition of survival of microorganism as well as
 CC microbistatic inhibition of growth. Thus the peptides are useful as
 CC therapeutic agents, disinfectants, food preservatives, or medicaments.
 CC The peptides are also useful for treating a patient suffering from
 CC bacterial, viral, fungal or other infection. The theta defensins have
 CC high antimicrobial activity and low haemolytic activity. The present
 CC sequence represents a Rhesus theta defensin analogue peptide.

XX Sequence 18 AA;

Query Match 73.3%; Score 85; DB 8; Length 18;

Best Local Similarity 75.0%; Pred. No. 0.014;

Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCIC 16

Db 1 GICRCYCGRGICRCIC 16

Search completed: October 26, 2004, 15:34:54

Job time : 116 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 15:39:33 ; Search time 87.5 Seconds
(without alignments)
66.602 Million cell updates/sec

Title: US-10-009-317A-32

Perfect score: 116
Sequence: 1 GFCRCICTRGFCRCICTR 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1364641 seqs, 323758627 residues

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------|
| 1 | 116 | 100.0 | 18 | 14 | US-10-060-102-28 |
| 2 | 116 | 100.0 | 18 | 15 | US-10-427-715-2 |
| 3 | 116 | 100.0 | 18 | 15 | US-10-427-715-14 |
| 4 | 116 | 100.0 | 18 | 15 | US-10-721-839-28 |
| 5 | 103 | 88.8 | 18 | 14 | US-10-060-102-32 |
| 6 | 103 | 88.8 | 18 | 15 | US-10-721-839-32 |
| 7 | 101 | 87.1 | 18 | 14 | US-10-060-102-30 |
| 8 | 101 | 87.1 | 18 | 14 | US-10-313-994-1 |
| 9 | 101 | 87.1 | 18 | 15 | US-10-427-715-1 |
| 10 | 101 | 87.1 | 18 | 15 | US-10-427-715-12 |
| 11 | 101 | 87.1 | 18 | 15 | US-10-427-715-13 |
| 12 | 101 | 87.1 | 18 | 15 | US-10-427-715-24 |
| 13 | 101 | 87.1 | 18 | 15 | US-10-721-839-30 |

| | | | | | | |
|----|-----|------|----|----|------------------|-------------------|
| 14 | 100 | 86.2 | 18 | 15 | US-10-427-715-29 | Sequence 29, Appl |
| 15 | 93 | 80.2 | 18 | 14 | US-10-141-645-5 | Sequence 5, Appl |
| 16 | 93 | 80.2 | 18 | 14 | US-10-141-645-6 | Sequence 6, Appl |
| 17 | 91 | 78.4 | 18 | 14 | US-10-313-994-9 | Sequence 9, Appl |
| 18 | 90 | 77.6 | 18 | 14 | US-10-060-102-27 | Sequence 27, Appl |
| 19 | 90 | 77.6 | 18 | 14 | US-10-141-645-1 | Sequence 1, Appl |
| 20 | 90 | 77.6 | 18 | 15 | US-10-721-839-27 | Sequence 27, Appl |
| 21 | 89 | 76.7 | 18 | 15 | US-10-427-715-23 | Sequence 23, Appl |
| 22 | 88 | 75.9 | 18 | 14 | US-10-060-102-31 | Sequence 31, Appl |
| 23 | 88 | 75.9 | 18 | 15 | US-10-721-839-31 | Sequence 31, Appl |
| 24 | 87 | 75.0 | 18 | 9 | US-09-917-340-53 | Sequence 53, Appl |
| 25 | 87 | 75.0 | 18 | 14 | US-10-141-645-2 | Sequence 2, Appl |
| 26 | 86 | 74.1 | 18 | 14 | US-10-060-102-29 | Sequence 29, Appl |
| 27 | 86 | 74.1 | 18 | 15 | US-10-427-715-3 | Sequence 3, Appl |
| 28 | 86 | 74.1 | 18 | 15 | US-10-427-715-15 | Sequence 15, Appl |
| 29 | 86 | 74.1 | 18 | 15 | US-10-427-715-16 | Sequence 16, Appl |
| 30 | 86 | 74.1 | 18 | 15 | US-10-721-839-29 | Sequence 29, Appl |
| 31 | 85 | 73.3 | 18 | 14 | US-10-141-645-3 | Sequence 3, Appl |
| 32 | 85 | 73.3 | 18 | 14 | US-10-141-645-4 | Sequence 4, Appl |
| 33 | 85 | 73.3 | 18 | 15 | US-10-427-715-30 | Sequence 30, Appl |
| 34 | 83 | 71.6 | 18 | 14 | US-10-141-645-7 | Sequence 7, Appl |
| 35 | 83 | 71.6 | 18 | 15 | US-10-427-715-19 | Sequence 19, Appl |
| 36 | 83 | 71.6 | 18 | 15 | US-10-427-715-20 | Sequence 20, Appl |
| 37 | 82 | 70.7 | 18 | 15 | US-10-427-715-17 | Sequence 17, Appl |
| 38 | 82 | 70.7 | 18 | 15 | US-10-427-715-31 | Sequence 31, Appl |
| 39 | 82 | 70.7 | 18 | 14 | US-10-141-645-8 | Sequence 8, Appl |
| 40 | 75 | 64.7 | 18 | 14 | US-10-141-645-9 | Sequence 9, Appl |
| 41 | 74 | 63.8 | 18 | 15 | US-10-427-715-28 | Sequence 28, Appl |
| 42 | 73 | 62.9 | 18 | 15 | US-10-427-715-37 | Sequence 37, Appl |
| 43 | 73 | 62.9 | 18 | 15 | US-10-427-715-38 | Sequence 38, Appl |
| 44 | 73 | 62.9 | 18 | 15 | US-10-427-715-39 | Sequence 39, Appl |
| 45 | 73 | 62.9 | 92 | 14 | US-10-313-994-21 | Sequence 21, Appl |

ALIGNMENTS

RESULT 1
US-10-060-102-28
; Sequence 28, Application US/10060102
; Publication No. US20030022829A1
; GENERAL INFORMATION:
; APPLICANT: MAURY, WENDY
; APPLICANT: STAPLETON, JACK
; APPLICANT: ROLLER, RICHARD
; APPLICANT: STINSKI, MARK
; APPLICANT: MCCRAY, PAUL B.
; APPLICANT: TACK, BRIAN
; TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIMATE THETA DEFENSINS AND MAMMALIAN
; FILE REFERENCE: IOWA:035US
; CURRENT APPLICATION NUMBER: US/10/060,102
; CURRENT FILING DATE: 2002-02-22
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/309,368
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,270
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 28
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-060-102-28

Query Match 100.0%; Score 116; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GFCRCICTRGFCRCICTR 18

USPA
Priority of current
app 11/06/04
18 May 2004

us-10-009-317a-32.rapb

Wed Oct 27 08:15:52 2004

```
Db 1 GFCRCICTRGFCRCICTR 18
|||||
RESULT 2
US-10-427-715-2
; Sequence 2, Application US/10427715
; Publication No. US20040014669A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tran, Dat Q.
; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
; TITLE OF INVENTION: Thereof, and Methods of Use
; FILE REFERENCE: 66778-302(UC5754)
; CURRENT APPLICATION NUMBER: US/10/427,715
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/377,071
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-10-427-715-2
Query Match 100.0%; Score 116; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
|||||
Db 1 GFCRCICTRGFCRCICTR 18
|||||

RESULT 3
US-10-427-715-14
; Sequence 14, Application US/10427715
; Publication No. US20040014669A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tran, Dat Q.
; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
; TITLE OF INVENTION: Thereof, and Methods of Use
; FILE REFERENCE: 66778-302(UC5754)
; CURRENT APPLICATION NUMBER: US/10/427,715
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/377,071
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-427-715-14
Query Match 100.0%; Score 116; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
|||||
Db 1 GFCRCICTRGFCRCICTR 18
|||||

RESULT 4
US-10-721-839-28
; Sequence 28, Application US/10721839
; Publication No. US20040086535A1
; GENERAL INFORMATION:
; APPLICANT: MAURY, WENDY
; APPLICANT: STAPLETON, JACK
; APPLICANT: ROLLER, RICHARD
; APPLICANT: STINSKI, MARK
; APPLICANT: MCCRAY, PAUL B.
; APPLICANT: TACK, BRIAN
; TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIMATE THETA DEFENSINS AND MAMMALI
; FILE REFERENCE: IOWA:035US
; CURRENT APPLICATION NUMBER: US/10/721,839
; CURRENT FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: US/10/060,102
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/309,368
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/265,270
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-060-102-32
Query Match 88.8%; Score 103; DB 14; Length 18;
Best Local Similarity 88.9%; Pred. No. 2.9e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
|||||
Db 1 GFCRCICTRGFCRCICTR 18
|||||

RESULT 5
US-10-060-102-32
; Sequence 32, Application US/10060102
; Publication No. US20030022829A1
; GENERAL INFORMATION:
; APPLICANT: MAURY, WENDY
; APPLICANT: STAPLETON, JACK
; APPLICANT: ROLLER, RICHARD
; APPLICANT: STINSKI, MARK
; APPLICANT: MCCRAY, PAUL B.
; APPLICANT: TACK, BRIAN
; TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIMATE THETA DEFENSINS AND MAMMALI
; FILE REFERENCE: IOWA:035US
; CURRENT APPLICATION NUMBER: US/10/060,102
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/309,368
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/265,270
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-060-102-32
Query Match 100.0%; Score 116; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
|||||
Db 1 GFCRCICTRGFCRCICTR 18
|||||
```

6008
claim #1
3/2/04

Db 1 GICRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | | | | | |

RESULT 6

US-10-721-839-32
; Sequence 32, Application US/10721839
; Publication No. US20040086535A1
; GENERAL INFORMATION:
; APPLICANT: MAURY, WENDY
; APPLICANT: STAPLETON, JACK
; APPLICANT: ROLLER, RICHARD
; APPLICANT: STINSKI, MARK
; APPLICANT: MCCRAY, PAUL B.
; APPLICANT: TACK, BRIAN
; TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIMATE THETA DEFENSINS AND MAMMALIAN CATHELICIDINS
; FILE REFERENCE: IOWA:035US
; CURRENT APPLICATION NUMBER: US/10/721,839
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: US/10/060,102
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/309,368
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/265,270
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 32
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-721-839-32

Query Match 88.8%; Score 103; DB 15; Length 18;
Best Local Similarity 88.9%; Pred. No. 2.9e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GICRCICTRGFCRCICTR 18

RESULT 7

US-10-060-102-30
; Sequence 30, Application US/10060102
; Publication No. US2003002829A1
; GENERAL INFORMATION:
; APPLICANT: MAURY, WENDY
; APPLICANT: STAPLETON, JACK
; APPLICANT: ROLLER, RICHARD
; APPLICANT: STINSKI, MARK
; APPLICANT: MCCRAY, PAUL B.
; APPLICANT: TACK, BRIAN
; TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIMATE THETA DEFENSINS AND MAMMALIAN CATHELICIDINS
; FILE REFERENCE: IOWA:035US
; CURRENT APPLICATION NUMBER: US/10/060,102
; CURRENT FILING DATE: 2002-02-22
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/309,368
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/265,270
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 30
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-060-102-30

Query Match 87.1%; Score 101; DB 14; Length 18;
Best Local Similarity 83.3%; Pred. No. 4.8e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GFCRCICTRGFCRCICTR 18

RESULT 8

US-10-313-994-1
; Sequence 1, Application US/10313994
; Publication No. US20030162718A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/10/313,994
; CURRENT FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-10-313-994-1

Query Match 87.1%; Score 101; DB 14; Length 18;
Best Local Similarity 83.3%; Pred. No. 4.8e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GFCRCICTRGFCRCICTR 18

RESULT 9

US-10-427-715-1
; Sequence 1, Application US/10427715
; Publication No. US20040014669A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Iran, Dat Q.
; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
; TITLE OF INVENTION: Thereof, and Methods of Use
; FILE REFERENCE: 66778-302(UC5754)
; CURRENT APPLICATION NUMBER: US/10/427,715
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/377,071
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-10-427-715-1

Query Match 87.1%; Score 101; DB 15; Length 18;
Best Local Similarity 83.3%; Pred. No. 4.8e-05;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | | | | | |

Db 1 GFCRCCLCRGVCRICICTR 18

RESULT 10

US-10-427-715-12
 ; Sequence 12, Application US/10427715
 ; Publication No. US20040014669A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Selsted, Michael E.
 ; APPLICANT: Tran, Dat Q.
 ; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
 ; TITLE OF INVENTION: Thereof, and Methods of Use
 ; FILE REFERENCE: 66778-302(UC5754)
 ; CURRENT APPLICATION NUMBER: US/10/427,715
 ; CURRENT FILING DATE: 2003-04-30
 ; PRIOR APPLICATION NUMBER: US 60/377,071
 ; PRIOR FILING DATE: 2002-04-30
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: synthetic construct
 US-10-427-715-12

Query Match 87.1%; Score 101; DB 15; Length 18;
 Best Local Similarity 83.3%; Pred. No. 4.8e-05;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICICTR 18

Db 1 GFCRCCLCRGVCRICICTR 18

RESULT 11

US-10-427-715-13
 ; Sequence 13, Application US/10427715
 ; Publication No. US20040014669A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Selsted, Michael E.
 ; APPLICANT: Tran, Dat Q.
 ; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
 ; TITLE OF INVENTION: Thereof, and Methods of Use
 ; FILE REFERENCE: 66778-302(UC5754)
 ; CURRENT APPLICATION NUMBER: US/10/427,715
 ; CURRENT FILING DATE: 2003-04-30
 ; PRIOR APPLICATION NUMBER: US 60/377,071
 ; PRIOR FILING DATE: 2002-04-30
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: synthetic construct
 ; NAME/KEY: AMIDATION
 ; LOCATION: 18
 ; OTHER INFORMATION: at the C terminus
 US-10-427-715-13

Query Match 87.1%; Score 101; DB 15; Length 18;
 Best Local Similarity 83.3%; Pred. No. 4.8e-05;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICICTR 18

Db 1 GFCRCCLCRGVCRICICTR 18

RESULT 12

US-10-427-715-24
 ; Sequence 24, Application US/10427715
 ; Publication No. US20040014669A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Selsted, Michael E.
 ; APPLICANT: Tran, Dat Q.
 ; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
 ; TITLE OF INVENTION: Thereof, and Methods of Use
 ; FILE REFERENCE: 66778-302(UC5754)
 ; CURRENT APPLICATION NUMBER: US/10/427,715
 ; CURRENT FILING DATE: 2003-04-30
 ; PRIOR APPLICATION NUMBER: US 60/377,071
 ; PRIOR FILING DATE: 2002-04-30
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 24
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: synthetic construct
 US-10-427-715-24

Query Match 87.1%; Score 101; DB 15; Length 18;
 Best Local Similarity 88.9%; Pred. No. 4.8e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICICTR 18

Db 1 GFCRCRCRCRGVCRICICTR 18

RESULT 13

US-10-721-839-30
 ; Sequence 30, Application US/10721839
 ; Publication No. US20040086535A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MAURY, WENDY
 ; APPLICANT: STAPLETON, JACK
 ; APPLICANT: ROLLER, RICHARD
 ; APPLICANT: STINSKI, MARK
 ; APPLICANT: MCCRAY, PAUL B.
 ; APPLICANT: TACK, BRIAN
 ; TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIMATE THETA DEFENSINS AND MAMMALIAN
 ; TITLE OF INVENTION: CATHELICIDINS
 ; FILE REFERENCE: IOWA:035US
 ; CURRENT APPLICATION NUMBER: US/10/721,839
 ; CURRENT FILING DATE: 2003-11-25
 ; PRIOR APPLICATION NUMBER: US/10/060,102
 ; PRIOR FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: 60/309,368
 ; PRIOR FILING DATE: 2001-08-01
 ; PRIOR APPLICATION NUMBER: 60/265,270
 ; PRIOR FILING DATE: 2001-01-30
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 30
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: Peptide
 US-10-721-839-30

Query Match 87.1%; Score 101; DB 15; Length 18;
 Best Local Similarity 83.3%; Pred. No. 4.8e-05;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICICTR 18

Db 1 GFCRCRCRGVCRICICTR 18

RESULT 16

APPLICANT: MAURY, WENDY
APPLICANT: STAPLETON, JACK
APPLICANT: ROLLER, RICHARD
APPLICANT: STINSKI, MARK
APPLICANT: MCCRAY, PAUL B.
APPLICANT: TACK, BRIAN
TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIVATE THETA DEFENSINS AND MAMMALI
TITLE OF INVENTION: CATHELICIDINS
FILE REFERENCE: IOWA-035US
CURRENT APPLICATION NUMBER: US/10/060,102
CURRENT FILING DATE: 2002-02-22
PRIOR APPLICATION NUMBER: 60/309,368
PRIOR FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/265,270
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 27
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-060-102-27

Query Match 77.6%; Score 90; DB 14; Length 18;
Best Local Similarity 77.8%; Pred. No. 0.00081;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | |
Db 1 GICRCICGRGICRCICGR 18

RESULT 19

US-10-141-645-1
Sequence 1, Application US/10141645
Publication No. US20030144184A1
GENERAL INFORMATION:
APPLICANT: Robert Lehrer
APPLICANT: Alan Waring
APPLICANT: Alexander Cole
APPLICANT: Teresa Hong
TITLE OF INVENTION: Retrocyclins - Antiviral and
TITLE OF INVENTION: Antimicrobial Peptides
FILE REFERENCE: UCLA-001CIP
CURRENT APPLICATION NUMBER: US/10/141,645
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: 60/284,855
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: Unassigned
PRIOR FILING DATE: 2002-04-18
NUMBER OF SEQ ID NOS: 125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 18
TYPE: PRT
ORGANISM: Homo sapiens
US-10-141-645-1

Query Match 77.6%; Score 90; DB 14; Length 18;
Best Local Similarity 77.8%; Pred. No. 0.00081;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | |
Db 1 GICRCICGRGICRCICGR 18

RESULT 20

US-10-721-839-27
Sequence 27, Application US/10721839

Publication No. US20040086535A1
GENERAL INFORMATION:
APPLICANT: MAURY, WENDY
APPLICANT: STAPLETON, JACK
APPLICANT: ROLLER, RICHARD
APPLICANT: STINSKI, MARK
APPLICANT: MCCRAY, PAUL B.
APPLICANT: TACK, BRIAN
TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIMATE THETA DEFENSINS AND MAMMALI
TITLE OF INVENTION: CATHELICIDINS
FILE REFERENCE: IOWA-035US
CURRENT APPLICATION NUMBER: US/10/721,839
CURRENT FILING DATE: 2003-11-25
PRIOR APPLICATION NUMBER: US/10/060,102
PRIOR FILING DATE: 2002-02-22
PRIOR APPLICATION NUMBER: 60/309,368
PRIOR FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/265,270
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 27
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-721-839-27

Query Match 77.6%; Score 90; DB 15; Length 18;
Best Local Similarity 77.8%; Pred. No. 0.00081;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | |
Db 1 GICRCICGRGICRCICGR 18

RESULT 21

US-10-427-715-23
Sequence 23, Application US/10427715
Publication No. US20040014669A1
GENERAL INFORMATION:
APPLICANT: Selsted, Michael E.
APPLICANT: Tran, Dat Q.
TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
TITLE OF INVENTION: Therest, and Methods of Use
FILE REFERENCE: 66778-302(UC5754)
CURRENT APPLICATION NUMBER: US/10/427,715
CURRENT FILING DATE: 2003-04-30
PRIOR APPLICATION NUMBER: US 60/377,071
PRIOR FILING DATE: 2002-04-30
NUMBER OF SEQ ID NOS: 41
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic construct
US-10-427-715-23

Query Match 76.7%; Score 89; DB 15; Length 18;
Best Local Similarity 77.8%; Pred. No. 0.0011;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | |
Db 1 GFCRCICGRGICRCICGR 18

RESULT 22

US-10-060-102-31
; Sequence 31, Application US/10060102
; Publication No. US20030022829A1
; GENERAL INFORMATION:
; APPLICANT: MAURY, WENDY
; APPLICANT: STAPLETON, JACK
; APPLICANT: ROLLER, RICHARD
; APPLICANT: STINSKI, MARK
; APPLICANT: MCCRAY, PAUL B.
; APPLICANT: TACK, BRIAN
; TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIMATE THETA DEFENSINS AND MAMMALI
; FILE REFERENCE: IOWA:035US
; CURRENT APPLICATION NUMBER: US/10/060,102
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/309,368
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/265,270
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-060-102-31

Query Match 75.9%; Score 88; DB 14; Length 18;
Best Local Similarity 72.2%; Pred. No. 0.0014;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GFCRCICRGFCRCICTR 18
| | | | | | | | | | | | | | | | | | | |
Db 1 GICRCLCRGVCRICGR 18

RESULT 23
US-10-721-839-31
; Sequence 31, Application US/10721839
; Publication No. US2004008635A1
; GENERAL INFORMATION:
; APPLICANT: MAURY, WENDY
; APPLICANT: STAPLETON, JACK
; APPLICANT: ROLLER, RICHARD
; APPLICANT: STINSKI, MARK
; APPLICANT: MCCRAY, PAUL B.
; APPLICANT: TACK, BRIAN
; TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIMATE THETA DEFENSINS AND MAMMALI
; FILE REFERENCE: IOWA:035US
; CURRENT APPLICATION NUMBER: US/10/721,839
; CURRENT FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: US/10/060,102
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/309,368
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/265,270
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-721-839-31

Query Match 75.9%; Score 88; DB 15; Length 18;

Best Local Similarity 72.2%; Pred. No. 0.0014;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 1 GFCRCICRGFCRCICTR 18
| | | | | | | | | | | | | | | | | | | |
Db 1 GICRCLCRGVCRICGR 18
RESULT 24
US-09-917-340-53
; Sequence 53, Application US/09917340
; Patent No. US20020090369A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Christopher J.
; APPLICANT: McAnulty, Jonathan F.
; APPLICANT: Reid, Ted W.
; TITLE OF INVENTION: Transplant Media
; FILE REFERENCE: TPLANT-06468
; CURRENT APPLICATION NUMBER: US/09/917,340
; CURRENT FILING DATE: 2001-07-29
; PRIOR APPLICATION NUMBER: 60/221,632
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/249,602
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/290,932
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-917-340-53

Query Match 75.0%; Score 87; DB 9; Length 18;
Best Local Similarity 86.7%; Pred. No. 0.0018;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | | | | | |
Db 1 RCICTRGFCRCICTR 15

RESULT 25
US-10-141-645-2
; Sequence 2, Application US/10141645
; Publication No. US20030144184A1
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; TITLE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-COICIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic variant
US-10-141-645-2

Query Match 75.0%; Score 87; DB 14; Length 18;
Best Local Similarity 72.2%; Pred. No. 0.0018;

```
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 GFCRCICTRGFCRCICTR 18
| | | | | : | | | | |
DB 1 GVCRCLCRRGVCRCICRR 18
| | | | | : | | | | |

RESULT 26
US-10-060-102-29
; Sequence 29, Application US/10060102
; Publication No. US20030022829A1
; GENERAL INFORMATION:
; APPLICANT: MAURY, WENDY
; APPLICANT: STAPLETON, JACK
; APPLICANT: ROLLER, RICHARD
; APPLICANT: STINSKI, MARK
; APPLICANT: MCCRAY, PAUL B.
; APPLICANT: TACK, BRIAN
; TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIMATE THETA DEFENSINS AND ANALOGS
; FILE REFERENCE: IOWA-035US
; CURRENT APPLICATION NUMBER: US/10/060,102
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/309,368
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/265,270
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 29
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-060-102-29

Query Match 74.1%; Score 86; DB 14; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0023;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 GFCRCICTRGFCRCICTR 18
| | | | | : | | | | |
DB 1 GVCRCLCRRGVCRCICRR 18
| | | | | : | | | | |

RESULT 27
US-10-427-715-3
; Sequence 3, Application US/10427715
; Publication No. US20040014669A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tran, Dat Q.
; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
; FILE REFERENCE: 66778-302(UC5754)
; CURRENT APPLICATION NUMBER: US/10/427,715
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/377,071
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Macaca mulatta
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-427-715-3

Query Match 74.1%; Score 86; DB 15; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0023;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 GFCRCICTRGFCRCICTR 18
| | | | | : | | | | |
DB 1 GVCRCLCRRGVCRCICRR 18
| | | | | : | | | | |

RESULT 28
US-10-427-715-15
; Sequence 15, Application US/10427715
; Publication No. US20040014669A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tran, Dat Q.
; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
; FILE REFERENCE: 66778-302(UC5754)
; CURRENT APPLICATION NUMBER: US/10/427,715
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/377,071
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-427-715-15

Query Match 74.1%; Score 86; DB 15; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0023;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 GFCRCICTRGFCRCICTR 18
| | | | | : | | | | |
DB 1 GVCRCLCRRGVCRCICRR 18
| | | | | : | | | | |

RESULT 29
US-10-427-715-16
; Sequence 16, Application US/10427715
; Publication No. US20040014669A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tran, Dat Q.
; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
; FILE REFERENCE: 66778-302(UC5754)
; CURRENT APPLICATION NUMBER: US/10/427,715
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/377,071
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-427-715-16

Query Match 74.1%; Score 86; DB 15; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0023;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 GFCRCICTRGFCRCICTR 18
| | | | | : | | | | |
DB 1 GVCRCLCRRGVCRCICRR 18
| | | | | : | | | | |
```

```
RESULT 30
US-10-721-839-29
; Sequence 29, Application US/10721839
; Publication No. US20040086535A1
; GENERAL INFORMATION:
; APPLICANT: MAURY, WENDY
; APPLICANT: STAPLETON, JACK
; APPLICANT: ROLLER, RICHARD
; APPLICANT: STINSKI, MARK
; APPLICANT: MCCRAY, PAUL B.
; APPLICANT: TACK, BRIAN
; TITLE OF INVENTION: NOVEL ANTIVIRAL ACTIVITIES OF PRIVATE THETA DEFENSINS AND MAMMALI
; FILE OF INVENTION: CATHELICIDINS
; FILE REFERENCE: IOWA:0350S
; CURRENT APPLICATION NUMBER: US/10/721,839
; CURRENT FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: US/10/060,102
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/309,368
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/265,270
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-721-839-29

Query Match 74.1%; Score 86; DB 15; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0023;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | |
DB 1 GVCRLCRRGVCRCICRR 18
| | | | | | | | | | | | | | | |

RESULT 31
US-10-141-645-3
; Sequence 3, Application US/10141645
; Publication No. US20030144184A1
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic variant
US-10-141-645-3

Query Match 73.3%; Score 85; DB 14; Length 18;
Best Local Similarity 72.2%; Pred. No. 0.0029;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | |
DB 1 GVCRCICRGVCRCICGR 18
| | | | | | | | | | | | | | | |

RESULT 32
US-10-141-645-4
; Sequence 4, Application US/10141645
; Publication No. US20030144184A1
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic variant
US-10-141-645-4

Query Match 73.3%; Score 85; DB 14; Length 18;
Best Local Similarity 72.2%; Pred. No. 0.0029;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
| | | | | | | | | | | | | | | |
DB 1 GVCRCICRGVCRCICGR 18
| | | | | | | | | | | | | | | |

RESULT 33
US-10-427-715-30
; Sequence 30, Application US/10427715
; Publication No. US20040014669A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tran, Dat Q.
; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
; FILE OF INVENTION: Thereof, and Methods of Use
; FILE REFERENCE: 66778-302(UC5754)
; CURRENT APPLICATION NUMBER: US/10/427,715
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/377,071
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-427-715-30

Query Match 73.3%; Score 85; DB 15; Length 18;
Best Local Similarity 75.0%; Pred. No. 0.0029;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

QY 1 GFCRCICTRGFCRCIC 16
Db 1 GICRCLCRGVGCRCIC 16

RESULT 34
US-10-141-645-7
; Sequence 7, Application US/10141645
; Publication No. US2003014184A1
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; TITLE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic variant
US-10-141-645-7

Query Match 71.6%; Score 83; DB 14; Length 18;
Best Local Similarity 72.2%; Pred. No. 0.0049;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCIC 18
Db 1 GICVCICGRCICICGR 18

RESULT 35
US-10-427-715-19
; Sequence 19, Application US/10427715
; Publication No. US20040014669A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tran, Dat Q.
; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
; TITLE OF INVENTION: Thereof, and Methods of Use
; FILE REFERENCE: 66778-302(UC5754)
; CURRENT APPLICATION NUMBER: US/10/427,715
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/377,071
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-427-715-19

Query Match 71.6%; Score 83; DB 15; Length 18;
Best Local Similarity 72.2%; Pred. No. 0.0049;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCIC 18
Db 1 GFCRALCRGVGCRAICR 18

RESULT 36
US-10-427-715-20
; Sequence 20, Application US/10427715
; Publication No. US20040014669A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tran, Dat Q.
; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
; TITLE OF INVENTION: Thereof, and Methods of Use
; FILE REFERENCE: 66778-302(UC5754)
; CURRENT APPLICATION NUMBER: US/10/427,715
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/377,071
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-427-715-20

Query Match 71.6%; Score 83; DB 15; Length 18;
Best Local Similarity 72.2%; Pred. No. 0.0049;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCIC 18
Db 1 GFCALCRGVGCRAICR 18

RESULT 37
US-10-427-715-17
; Sequence 17, Application US/10427715
; Publication No. US20040014669A1
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tran, Dat Q.
; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
; TITLE OF INVENTION: Thereof, and Methods of Use
; FILE REFERENCE: 66778-302(UC5754)
; CURRENT APPLICATION NUMBER: US/10/427,715
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/377,071
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
; NAME/KEY: AVIDATION
; LOCATION: 18
; OTHER INFORMATION: at the C terminus
US-10-427-715-17

Query Match 70.7%; Score 82; DB 15; Length 18;
Best Local Similarity 68.8%; Pred. No. 0.0064;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCIC 16
Db 2 GVCRLCRGVGCRLC 17

RESULT 38

```
US-10-427-715-31
; Sequence 31, Application US/10427715
; Publication No. US20040014669A1
; GENERAL INFORMATION:
; APPLICANT: Tran, Dat Q.
; APPLICANT: Selsted, Michael E.
; TITLE OF INVENTION: Antimicrobial Theta Defensins, Analogs
; FILE REFERENCE: 86778-302(UC5754)
; CURRENT APPLICATION NUMBER: US/10/427,715
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/377,071
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-427-715-31

Query Match          70.7%; Score 82; DB 15; Length 18;
Best Local Similarity 75.0%; Pred. No. 0.0064;
Matches 12; Conservative 0; Mismatches 4; Indels 4; Gaps 0;

QY 1 GFCRCICTRGFCRCIC 16
DB 1 GICRCICVLGICRCIC 16

RESULT 39
US-10-141-645-8
; Sequence 8, Application US/10141645
; Publication No. US20030144184A1
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic variant
US-10-141-645-8

Query Match          64.7%; Score 75; DB 14; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.039;
Matches 12; Conservative 0; Mismatches 6; Indels 6; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
DB 1 GICICICGVGICRCICGR 18

Search completed: October 26, 2004, 15:55:49
Job time : 88.5 secs

US-10-141-645-9
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic variant
US-10-141-645-9

Query Match          64.7%; Score 75; DB 14; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.039;
Matches 12; Conservative 0; Mismatches 6; Indels 6; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
DB 1 GICICICGVGICRCICGR 18

RESULT 40
US-10-141-645-9
; Sequence 9, Application US/10141645
; Publication No. US20030144184A1
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 15:13:56 ; Search time 128.5 Seconds
(without alignments)
80.597 Million cell updates/sec

Title: US-10-009-317A-32

Perfect score: 116

Sequence: 1 GFCRCITRGFCRCICTR 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------|
| 1 | 64 | 55.2 | 76 | P82270 | macaca mula |
| 2 | 64 | 55.2 | 76 | Q9TU01 | macaca mula |
| 3 | 59 | 50.9 | 168 | Q6P8T4 | mus musculus |
| 4 | 59 | 50.9 | 168 | Q8CH20 | mus musculus |
| 5 | 59 | 50.9 | 168 | Q9D912 | mus musculus |
| 6 | 59 | 50.9 | 168 | AAH61079 | mus musculus |
| 7 | 59 | 50.9 | 173 | Q9D4K2 | mus musculus |
| 8 | 59 | 50.9 | 274 | Q949G1 | oryza sativa |
| 9 | 56.5 | 48.7 | 163 | Q8VJ20 | mycobacteri |
| 10 | 56 | 48.3 | 937 | Q8BLJ1 | ciona intes |
| 11 | 54.5 | 47.0 | 307 | Q9S282 | plasmodium |
| 12 | 54 | 46.6 | 290 | Y070 | autographa |
| 13 | 54 | 46.6 | 991 | Q7SEQ2 | neurospora |
| 14 | 53.5 | 46.1 | 174 | Q9NRB6 | homo sapien |
| 15 | 53 | 45.7 | 164 | Q22048 | caenorhabdi |
| 16 | 53 | 45.7 | 166 | Q95QY1 | caenorhabdi |
| 17 | 53 | 45.7 | 188 | Q18238 | caenorhabdi |
| 18 | 53 | 45.7 | 197 | Q17641 | caenorhabdi |
| 19 | 53 | 45.7 | 602 | Q7SNZ5 | chlamydomon |
| 20 | 53 | 45.7 | 602 | BAD13491 | chlamydomon |
| 21 | 53 | 45.7 | 602 | BAD13492 | chlamydomon |
| 22 | 53 | 45.7 | 1365 | Q7SN88 | homo sapien |
| 23 | 53 | 45.7 | 1365 | BAD16738 | homo sapien |
| 24 | 53 | 45.7 | 2871 | Q7SN87 | homo sapien |
| 25 | 53 | 45.7 | 2871 | BAD16739 | homo sapien |
| 26 | 52.5 | 45.3 | 129 | Q7FU70 | anopheles g |
| 27 | 52.5 | 45.3 | 273 | Q94UZ6 | leishmania |
| 28 | 52.5 | 45.3 | 512 | Q6RY99 | rattus norv |
| 29 | 52.5 | 45.3 | 512 | AAH24072 | rattus norv |
| 30 | 52 | 44.8 | 66 | Q9PYQ3 | xestia c-ni |
| 31 | 52 | 44.8 | 146 | Q7QBV4 | anopheles g |

| | | | | | | |
|----|------|------|------|---|------------|-------------------|
| 32 | 52 | 44.8 | 161 | 2 | Q8MZ55 | Osm255 drosophila |
| 33 | 52 | 44.8 | 1823 | 2 | Q7PRP5 | anopheles g |
| 34 | 51.5 | 44.4 | 190 | 2 | Q9UI23 | homo sapien |
| 35 | 51 | 44.0 | 59 | 2 | Q6IG37 | drosophila |
| 36 | 51 | 44.0 | 72 | 1 | IBB_PHAU | phaseolus a |
| 37 | 51 | 44.0 | 201 | 2 | Q6ZQ52 | homo sapien |
| 38 | 51 | 44.0 | 201 | 2 | BAC87611 | homo sapi |
| 39 | 51 | 44.0 | 326 | 2 | Q91ZHE | meriones un |
| 40 | 51 | 44.0 | 336 | 2 | Q7PVN9 | anopheles g |
| 41 | 51 | 44.0 | 415 | 1 | VEGC MOUSE | mus musculu |
| 42 | 51 | 44.0 | 415 | 2 | Q91ZE3 | rattus norv |
| 43 | 51 | 44.0 | 415 | 2 | BAC33172 | mus muscu |
| 44 | 51 | 44.0 | 419 | 1 | VEGC_HUMAN | homo sapien |
| 45 | 51 | 44.0 | 419 | 2 | AAH63685 | homo sapi |

ALIGNMENTS

RESULT 1

P82270 P82270 PRELIMINARY; PRT; 76 AA.
AC P82270;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Theta defensin-1, subunit A precursor (RTD-1).
GN Name:RTD1A;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 65-73, AND MASS SPECTROMETRY.
RC TISSUE=Bone marrow, and Leukocyte;
RX MEDLINE=99453140; PubMed=10521339;
RA Tang Y.-O., Yuan J., Oesapay G., Oesapay K., Tran D., Miller C.J.,
RA Ouellette A.J., Selsted M.E.;
RT "A Cyclic Antimicrobial Peptide Produced in Primate Leukocytes by the
RT Ligation of Two Truncated alpha-Defensins.";
RL Science 286:498-502(1999).
CC -!- FUNCTION: Microbicidal activity against Gram-positive bacteria
CC S.aureus and L.monocytogenes, Gram-negative bacteria S.typhimurium
CC and E.coli ML35 and fungi C.albicans and C.neoformans in vitro.
CC -!- SUBUNIT: Heterodimer of subunit A and subunit B linked by a
CC disulfid bond at position 66 forming a cyclic RTD-1.
CC -!- TISSUE SPECIFICITY: Bone marrow; promyelocytes, myelocytes and
CC mature neutrophils and monocytes.
CC -!- DEVELOPMENTAL STAGE: Expression and peptide accumulation starts
CC early during granulocyte myelopoiesis.
CC -!- MASS SPECTROMETRY: MW=2082.0; METHOD=VALDI.
CC -!- MISCELLANEOUS: The determined pI of this protein is greater than
CC 12.
CC -!- SIMILARITY: Belongs to the corticostatin/defensin family.
CC EMBL; AF191100; AAF04389.1; -.
CC EMBL; AF191102; AAF04391.1; -.
CC PIR; A59089; A59089.
CC GO; GO:0050832; P:defense response to fungi; IEA.
CC GO; GO:0006805; P:xenobiotic metabolism; IEA.
CC InterPro; IPR002366; Defensin_propep.
CC Pram; PF00879; Defensin_propep; 1.
CC KW Antibiotic; Defensin; Fungicide; Signal.
CC SIGNAL 1 20 POTENTIAL.
CC PROPEP 21 64
CC CHAIN 65 73 THETA DEFENSIN-1, SUBUNIT A.
CC PROPEP 74 76 REMOVED IN MATURE FORM.
CC DISULFID 68 73
CC SEQUENCE 76 AA; 8242 MW; BEA207932A030590 CRC64;

Query Match 55.2%; Score 64; DB 2; Length 76;
Best Local Similarity 83.3%; Pred. No. 0.35;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCICTRGFCRCI 15
 RP ||||| :
 DB 65 RCICTRGFCRCI 76

RESULT 2

ID Q9TU01 PRELIMINARY; PRT; 76 AA.
 AC Q9TU01;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Demifensin 2.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 CX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhao C., Nguyen T., Lehrer R.I.,
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF184157; AAF07924.1; -
 DR GO: GO:0006952; P: defense response; IEA.
 DR InterPro: IPR002366; Defensein_propep.
 DR Pfam: PF00879; Defensein_propep; 1.
 SQ SEQUENCE 76 AA; 8212 MW; 1EB307932A031826 CRC64;

Query Match 55.2%; Score 64; DB 2; Length 76;
 Best Local Similarity 83.3%; Pred. No. 0.35;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCICTRGFCRCI 15
 RP ||||| :
 DB 65 RCICTRGFCRCI 76

RESULT 3

ID Q6P8T4 PRELIMINARY; PRT; 168 AA.
 AC Q6P8T4
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik K., Farmer A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Franco C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerk A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RP "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Strausberg R.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC061079; AAHG1079.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 168 AA; 18947 MW; B00FD3D27B9BF768 CRC64;

Query Match 50.9%; Score 59; DB 2; Length 168;
 Best Local Similarity 56.2%; Pred. No. 3;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 RCICTRGFCRCICTR 18
 RP ||||| :
 DB 66 CRCCCYCRCCRCRCSR 81

RESULT 4

ID Q8CH20 PRELIMINARY; PRT; 168 AA.
 AC Q8CH20;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Basic protein CMT1R3
 GN Name=4931420D14Rik; Synonyms=Ckl1r3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Testis;
 RA Xu X., Bai X., Silvius D., Escalier D., McFarland L., Xu P.-X.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF463502; AAQ15675.1; -
 DR MGD; MGI:1913992; 4931420D14Rik.
 SQ SEQUENCE 168 AA; 18957 MW; 8F30D3D27B9BF595 CRC64;

Query Match 50.9%; Score 59; DB 2; Length 168;
 Best Local Similarity 56.2%; Pred. No. 3;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 RCICTRGFCRCICTR 18
 RP ||||| :
 DB 66 CRCCCYCRCCRCRCSR 81

RESULT 5

ID Q9D9I2 PRELIMINARY; PRT; 168 AA.
 AC Q9D9I2;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
 DE library, clone:1700065I05 product:hypothetical Cysteine-rich region
 DE containing protein, full insert sequence.
 GN Name=4931420D14Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RL "High-efficiency full-length cDNA cloning.";
 RT Meth. Enzymol. 303:19-44(1999).
 RN [2]

RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:695-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA The FANTOM Consortium.
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA MEDLINE=20530913; PubMed=11076861;
 RA Konno H., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Haraoka T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AK016467; BAB30253.1;
 DR MGD; MGI:1913992; 4931420D14rik.
 KW Hypothetical protein.
 SQ SEQUENCE 173 AA; 19581 MW; 147B6F155AC29FDF CRC64;
 Query Match 50.9%; Score 59; DB 2; Length 173;
 Best Local Similarity 56.2%; Pred. No. 3.1;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 3 CRCICTRGFCRCICTR 18
 DB 66 CCCCCCCTCCGCCCR 81
 RESULT 8
 Q949G1 ID Q949G1 PRELIMINARY; PRT; 274 AA.
 AC Q949G1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein C15ERIPDM.
 GN Name=C15ERIPDM;
 OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriarthroideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21329048; PubMed=11435398;
 RA Mayer K., Murphy G., Tarchini R., Wambutt R., Volckaert G., Pohl T.,
 RA Duesterhoft A., Stiekema W., Entian K.D., Terry N., Lemcke K.,
 RA Haase D., Hall C.R., van Dodeweerd A.M., Tingey S.V., Mewes H.W.,
 RA Bevan M., Bancroft I.;
 RT "Conservation of microstructure between a sequenced region of the
 genome of rice and multiple segments of the genome of Arabidopsis
 thaliana";
 RL Genome Res. 11:1167-1174(2001).
 DR EMBL; AJ307662; CAC39030.1; -;
 DR Gramene; Q949G1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 274 AA; 28657 MW; AB547D9BD5470AE1 CRC64;
 Query Match 50.9%; Score 59; DB 2; Length 274;
 Best Local Similarity 66.7%; Pred. No. 4.5;
 Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 4 RCICTRGFCRCICTR 18
 DB 185 RCCCHGCCRCRATR 199
 RESULT 9
 Q9VJ20 ID Q9VJ20 PRELIMINARY; PRT; 163 AA.
 AC Q8VJ20;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein MT3454.
 GN OrderedLocustNames=MT3454;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA MEDLINE=22206494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
 RA Hickey E.K., Kolonay J.F., Nelson W.C., Unayam L.A., Emdolaeva M.D.,
 RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
 RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
 RA Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains";
 RL J. Bacteriol. 184:5479-5490(2002).
 DR EMBL; AF007152; AAK47795.1; -;
 DR TIGR; MT3454; -;
 KW Hypothetical protein.
 SQ SEQUENCE 163 AA; 18621 MW; B5E62AB951B2AC3C CRC64;
 Query Match 48.7%; Score 56.5; DB 2; Length 163;
 Best Local Similarity 60.0%; Pred. No. 6;
 Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;
 QY 3 CRCICTR-GFCRCIC 16
 DB 137 CRHVCTRSYGYCLVC 151
 RESULT 10
 Q9BLJ1 ID Q9BLJ1 PRELIMINARY; PRT; 937 AA.
 AC Q9BLJ1;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ci-METAL.
GN Name-Ci-metal;
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21347414; PubMed=11455433;
RA Nakayama A., Satou Y., Satou N.;
RT "Isolation and characterization of genes that are expressed during
RT Ciona intestinalis metamorphosis."
RL Dev. Genes Evol. 211:184-189(2001).
DR EMBL; AB041857; BAB40596.1; -
DR HSSP; P00743; 1CCF.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008009; F:chemokine activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007596; F:blood coagulation; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR000152; ASX_hydroxyl_S.
DR InterPro; IPR001811; Chemokine_IL6.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR009030; GROW_fac_recept.
DR InterPro; IPR001212; Somatomedin_B.
DR InterPro; IPR001491; Thrombomodulin.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF07645; EGF_Ca; 14.
DR PRINTS; PRO0022; SOMATOMEDIN.
DR PRINTS; PRO0907; THROMBOMODULN.
DR SMART; SM00179; EGF_CA; 14.
DR PROSITE; PS00010; ASX_HYDROXYL; 14.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 19.
DR PROSITE; PS00266; EGF_3; 19.
DR PROSITE; PS01187; EGF_CA; 12.
DR PROSITE; PS00524; SOMATOMEDIN_B; 1.
KW EGF-like domain.
SQ SEQUENCE 937 AA; 101043 MW; 8C67830C8E391D07 CRC64;

Query Match. 48.3%; Score 56; DB 2; Length 937;
Best Local Similarity 58.8%; Pred. No. 31;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 GFCICTGRCRCICT 17
DB 552 GSYRCICARGFGSGLCT 568

RESULT 11
ID 096282 PRELIMINARY; PRT; 307 AA.
AC 096282;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein PFB0950W.
GN Name=PFB0950W;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99021743; PubMed=9804551;
RA Gardner M.J., Tetelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shallow S., Mason T., Yu K., Fujii C., Pederson J.,

RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Perlea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Science 282:1126-1132(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan Y.-S., Nene V., Shallow S.J., Suh B., Peterson J., Agiusoli S.,
RA Perlea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramaniam L.G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE001428; AAC71979.2; -
DR PIR; E71602; E71602.
DR InterPro; IPR001368; TNFR_C6.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 307 AA; 35537 MW; B95A3DB354D4BE71 CRC64;

Query Match. 47.0%; Score 54.5; DB 2; Length 307;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 9; Conservative 1; Mismatches 2; Indels 3; Gaps 1;

QY 3 CRCICTGRCRCICT 17
DB 222 CSCICT---CTCICS 233

RESULT 12
ID Y070 NPVAC STANDARD; PRT; 290 AA.
AC P41470;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 34.4 kDa protein in LEF3-IAP2 intergenic region.
OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46015;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C6;
RX MEDLINE=94303173; PubMed=8030224;
RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Verber M., Possee R.D.;
RT "The complete DNA sequence of Autographa californica nuclear
RT polyhedrosis virus."
RL Virology 202:586-605(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. Usage by and for commercial
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; L22858; AAA66700.1; -
DR PIR; G72858; G72858.
DR InterPro; IPR000519; P_trefoil.
KW Hypothetical protein.
SQ SEQUENCE 290 AA; 34408 MW; CA78BA9C8B5AB937 CRC64;

Query Match. 46.6%; Score 54; DB 1; Length 290;

```

Best Local Similarity 40.7%; Pred. No. 21;
Matches 11; Conservative 2; Mismatches 4; Indels 10; Gaps 2;

QY 2 FRCICIT-----RGF-CRCICITR 18
    |||||
Db 218 FARCFCINTMQCPCPGQYKCEICIRR 244

RESULT 13
Q7SEQ2
ID Q7SEQ2 PRELIMINARY; PRT; 991 AA.
AC Q7SEQ2
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted protein.
DB Name=NCU02165.1;
GN Neurospora crassa.
OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Galvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,
RA Salitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gherre S.,
RA Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofora S., Rasmussen C., Mezenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmari S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Flammann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbold D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.,
RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 010-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABX01000029; EAA35288.1; -.
DR InterPro; IPR006209; EGF like.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS01186; EGF 2; 1.
SQ SEQUENCE 991 AA; 103024 MW; C51719B9F4D54A8E CRC64;

Query Match 46.6%; Score 54; DB 2; Length 991;
Best Local Similarity 56.2%; Pred. No. 59;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCIC 16
    |||||
Db 620 GFCSCICRNGFTGNC 635

RESULT 14
Q9NRB6 PRELIMINARY; PRT; 174 AA.
AC Q9NRB6
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mutant fibroblast growth factor receptor 3 (Fragment).
GN Name=RGFR3;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Chesi M., Brents L.A., Ely S.A., Bais C., Mesri E.A., Robbani D.,
RA Kuehl W.M., Bergsagel P.L.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238374; AAF97749.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_CODING
SQ SEQUENCE 174 AA; 17810 MW; BC9917E34470B9EA CRC64;

Query Match 45.1%; Score 53.5; DB 2; Length 174;
Best Local Similarity 47.1%; Pred. No. 15;
Matches 8; Conservative 2; Mismatches 4; Indels 3; Gaps 1;

QY 1 GFCRCICTRGFCRCICIT 17
    |||||
Db 75 GLCVVCVCV---CVCVCT 88

RESULT 15
Q22048 PRELIMINARY; PRT; 164 AA.
AC Q22048
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein T01B7.8.
GN Name=T01B7.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Sims M.A.;
RA MEDLINE=99069613; PubMed=9851916;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Sims M.A.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z66499; CAA91301.1; -.
DR PIR; T24272; T24272.
DR HSSP; P10968; 2CWG.
DR INACT; Q22048; -.CE03592.
DR WormPep; T01B7.8; CE03592.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR InterPro; IPR006081; Defensin alpha.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR001007; VWF_C.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; UNKNOWN_1.
DR PROSITE; PS00269; DEFENSIN; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01208; VWF_C_1; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 16499 MW; C002D48D36C9FCED CRC64;

Query Match 45.7%; Score 53; DB 2; Length 164;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICITR 18
    |||||
Db 80 GCGCGCCCRPRCCCCRR 97

RESULT 16
Q95QY1 PRELIMINARY; PRT; 166 AA.
ID Q95QY1
AC Q95QY1;

```

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C04G6.10.
GN ORFNames=C04G6.10;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Anderson K., Chissos S.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]
RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;
RA Nhan M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58760; AAK31463.1; -.
DR PIR; T15651; T15651.
DR HSSP; P10968; 2CWG.
DR IntAct; Q18238; -.
DR WormPep; C27A2.5; CE04105.
DR GO; GO:000576; C:extracellular; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR GO; GO:0009613; P:response to pest/pathogen/parasite; IEA.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR InterPro; IPR006081; Defensin_alpha.
DR InterPro; IPR001007; VWF_C.
DR PROSITE; PS00198; 4Fe4S_FERREDOXIN; UNKNOWN_1.
DR PROSITE; PS00269; DEFENSIN; 1.
DR PROSITE; PS01208; VWF_C_1; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 188 AA; 18878 MW; 0C5D0DC5CA8E0C4B CRC64;

Query Match 45.7%; Score 53; DB 2; Length 188;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 9; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
DB 81 GGCGCCCRPKCCCCRR 98

RESULT 18

ID Q17641 PRELIMINARY; PRT; 197 AA.
AC Q17641;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C04G6.7.
GN ORFNames=C04G6.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Anderson K., Chissos S.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58554; AAK68161.1; -.
DR HSSP; P10969; 1WGT.
DR WormPep; C04G6.7; CE27652.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR InterPro; IPR006081; Defensin_alpha.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR001007; VWF_C.
DR PROSITE; PS00198; 4Fe4S_FERREDOXIN; UNKNOWN_1.

Query Match 45.7%; Score 53; DB 2; Length 166;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GFCRCICTRGFCRCICTR 18
DB 81 GGCGCCCRPKCCCCRR 98

RESULT 17

ID Q18238 PRELIMINARY; PRT; 188 AA.
AC Q18238;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C27A2.5.
GN ORFNames=C27A2.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]
RP SEQUENCE FROM N.A.

DR PROSITE; PS00269; DEFENSIN; UNKNOWN_1.
 DR PROSITE; PS00222; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01208; VWFC_1; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 197 AA; 20596 MW; F55F9457BF9B9BAD CRC64;
 Query Match 45.7%; Score 53; DB 2; Length 197;
 Best Local Similarity 50.0%; Pred. No. 20;
 Matches 9; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 GFCRCICTRGFCRCICTR 18
 DB 82 GGGCCCCCPRCCCCCR 99
 RESULT 19
 Q75NZ5 PRELIMINARY; PRT; 602 AA.
 AC Q75NZ5
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 03-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Low-CO2 inducible Myb transcription factor LCR1.
 GN Names:LCR1;
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 OX NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yoshioka S., Taniguchi F., Miura K., Inoue T., Yamano T., Fukuzawa H.;
 RT "A Novel Myb Transcription Factor LCR1 Regulates the CO2-Responsive
 RT Gene Cahl Encoding A Periplasmic Carbonic Anhydrase in Chlamydomonas
 RT reinhardtii.";
 RL Plant Cell 0:0-0(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yoshioka S., Taniguchi F., Miura K., Inoue T., Yamano T., Fukuzawa H.;
 RT "A Novel Myb Transcription Factor LCR1 Regulates the CO2-Responsive
 RT Gene Cahl Encoding A Periplasmic Carbonic Anhydrase in Chlamydomonas
 RT reinhardtii.";
 RL Plant Cell 0:0-0(2004).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Asamizu E., Miura K., Kucho K., Inoue Y., Fukuzawa H., Ohshima K.,
 RA Nakamura Y., Tabata S.;
 RT "Generation of expressed sequence tags from low-CO2 and high-CO2
 RT adapted cells of Chlamydomonas reinhardtii.";
 RL DNA Res. 7:305-307(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Yoshioka S., Taniguchi F., Ohshima K., Fukuzawa H.;
 RT "Cis-acting elements and DNA-binding proteins involved in CO2-
 RT responsive transcriptional activation of Cahl encoding a periplasmic
 RT carbonic anhydrase in Chlamydomonas reinhardtii.";
 RL Plant Physiol. 133:783-793(2003).
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: Contains 1 Myb-like domain.
 DR EMBL; AB168090; BAD13492.1; -;
 DR EMBL; AB168089; BAD13491.1; -;
 DR InterPro; IPR009057; Homeodomain like.
 DR InterPro; IPR001005; Myb DNA binding.
 DR Pfam; PF00249; Myb DNA-binding; 1.
 DR SMART; SM00717; SANT; 1.
 DR PROSITE; PS50090; MYB_3; 1.
 KW Nuclear protein.
 SQ SEQUENCE 602 AA; 62736 MW; 3DD94253D5A792D0 CRC64;
 Query Match 45.7%; Score 53; DB 2; Length 602;
 Best Local Similarity 75.0%; Pred. No. 51;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 7 CTRGFCRCICTR 18
 DB 471 CTRCQCRCICCR 482
 RESULT 20
 Q75NZ5 PRELIMINARY; PRT; 602 AA.
 AC Q75NZ5
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 03-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Low-CO2 inducible Myb transcription factor LCR1.
 GN Names:LCR1;
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 OX NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yoshioka S., Taniguchi F., Miura K., Inoue T., Yamano T., Fukuzawa H.;
 RT "A Novel Myb Transcription Factor LCR1 Regulates the CO2-Responsive
 RT Gene Cahl Encoding A Periplasmic Carbonic Anhydrase in Chlamydomonas
 RT reinhardtii.";
 RL Plant Cell 0:0-0(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yoshioka S., Taniguchi F., Miura K., Inoue T., Yamano T., Fukuzawa H.;
 RT "A Novel Myb Transcription Factor LCR1 Regulates the CO2-Responsive
 RT Gene Cahl Encoding A Periplasmic Carbonic Anhydrase in Chlamydomonas
 RT reinhardtii.";
 RL Plant Cell 0:0-0(2004).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Asamizu E., Miura K., Kucho K., Inoue Y., Fukuzawa H., Ohshima K.,
 RA Nakamura Y., Tabata S.;
 RT "Generation of expressed sequence tags from low-CO2 and high-CO2
 RT adapted cells of Chlamydomonas reinhardtii.";
 RL DNA Res. 7:305-307(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Yoshioka S., Taniguchi F., Ohshima K., Fukuzawa H.;
 RT "Cis-acting elements and DNA-binding proteins involved in CO2-
 RT responsive transcriptional activation of Cahl encoding a periplasmic
 RT carbonic anhydrase in Chlamydomonas reinhardtii.";
 RL Plant Physiol. 133:783-793(2003).
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: Contains 1 Myb-like domain.
 DR EMBL; AB168090; BAD13492.1; -;
 DR EMBL; AB168089; BAD13491.1; -;
 DR InterPro; IPR009057; Homeodomain like.
 DR InterPro; IPR001005; Myb DNA binding.
 DR Pfam; PF00249; Myb DNA-binding; 1.
 DR SMART; SM00717; SANT; 1.
 DR PROSITE; PS50090; MYB_3; 1.
 KW Nuclear protein.
 SQ SEQUENCE 602 AA; 62736 MW; 3DD94253D5A792D0 CRC64;
 Query Match 45.7%; Score 53; DB 2; Length 602;
 Best Local Similarity 75.0%; Pred. No. 51;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 7 CTRGFCRCICTR 18
 DB 471 CTRCQCRCICCR 482

BAD13491 PRELIMINARY; PRT; 602 AA.
 ID BAD13491
 AC BAD13491
 DT 01-APR-2004 (TrEMBLrel. 27, Created)
 DT 01-APR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-APR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Low-CO2 inducible Myb transcription factor LCR1.
 GN LCR1.
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 OX NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yoshioka S., Taniguchi F., Miura K., Inoue T., Yamano T., Fukuzawa H.;
 RT "A Novel Myb Transcription Factor LCR1 Regulates the CO2-Responsive
 RT Gene Cahl Encoding A Periplasmic Carbonic Anhydrase in Chlamydomonas
 RT reinhardtii.";
 RL Plant Cell 0:0-0(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kucho K., Yoshioka S., Taniguchi F., Ohshima K., Fukuzawa H.;
 RT "Cis-acting elements and DNA-binding proteins involved in CO2-
 RT responsive transcriptional activation of Cahl encoding a periplasmic
 RT carbonic anhydrase in Chlamydomonas reinhardtii.";
 RL Plant Physiol. 133:783-793(2003).
 DR EMBL; AB168089; BAD13491.1; -;
 SQ SEQUENCE 602 AA; 62736 MW; 3DD94253D5A792D0 CRC64;
 Query Match 45.7%; Score 53; DB 2; Length 602;
 Best Local Similarity 75.0%; Pred. No. 51;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 7 CTRGFCRCICTR 18
 DB 471 CTRCQCRCICCR 482
 RESULT 21
 BAD13492 PRELIMINARY; PRT; 602 AA.
 ID BAD13492
 AC BAD13492
 DT 01-APR-2004 (TrEMBLrel. 27, Created)
 DT 01-APR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-APR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Low-CO2 inducible Myb transcription factor LCR1.
 GN LCR1.
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 OX NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yoshioka S., Taniguchi F., Miura K., Inoue T., Yamano T., Fukuzawa H.;
 RT "A Novel Myb Transcription Factor LCR1 Regulates the CO2-Responsive
 RT Gene Cahl Encoding A Periplasmic Carbonic Anhydrase in Chlamydomonas
 RT reinhardtii.";
 RL Plant Cell 0:0-0(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Asamizu E., Miura K., Kucho K., Inoue Y., Fukuzawa H., Ohshima K.,
 RA Nakamura Y., Tabata S.;
 RT "Generation of expressed sequence tags from low-CO2 and high-CO2
 RT adapted cells of Chlamydomonas reinhardtii.";
 RL DNA Res. 7:305-307(2000).
 DR EMBL; AB168090; BAD13492.1; -;
 SQ SEQUENCE 602 AA; 62736 MW; 3DD94253D5A792D0 CRC64;
 Query Match 45.7%; Score 53; DB 2; Length 602;
 Best Local Similarity 75.0%; Pred. No. 51;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 7 CTRGFCRCICTR 18
 DB 471 CTRCQCRCICCR 482

```
Db 471 CTRCQRCICR 482
|||||
Query Match 45.7%; Score 53; DB 2; Length 1365;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RT "3 Novel mutations of Fibrillin-1 and 10 single nucleotide
RT polymorphisms of Fibrillin-3 in Marfan syndrome patients.";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB177802; BAD16738.1; -.
SQ SEQUENCE 1365 AA; 147448 MW; ED83FB79C1B2BD48 CRC64;

Query Match 45.7%; Score 53; DB 2; Length 1365;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 RCICTRGFCRCICTR 18
|||||
Db 464 RCITPGSCRCCECNK 478

RESULT 24
Q75N87 PRELIMINARY; PRT; 2871 AA.
AC Q75N87;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Fibrillin 1.
GN Name=FN1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueda T., Takahashi T., Eto S., Sato T., Xu G., Toki T., Yonesaka S.,
RA Ito E.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
DR EMBL; AB177803; BAD16739.1; -.
DR InterPro; IPR000152; Asx_Hydroxyl_S.
DR InterPro; IPR008985; ConA_like_1ec_gi.
DR InterPro; IPR007042; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002212; Fibril-assoc.
DR InterPro; IPR011398; Fibrillin.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00008; EGF; 27.
DR Pfam; PF07645; EGF_CA; 42.
DR Pfam; PF00683; TB_5.
DR SMART; SMART0181; EGF; 47.
DR SMART; SMART0179; EGF_CA; 44.
DR PROSITE; PS00010; ASX_HYDROXYL; 42.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 38.
DR PROSITE; PS00026; EGF_3; 39.
DR PROSITE; PS01187; EGF_CA; 42.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW EGF-like domain.
SQ SEQUENCE 2871 AA; 312219 MW; C3D044A80321E082 CRC64;

Query Match 45.7%; Score 53; DB 2; Length 2871;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 RCICTRGFCRCICTR 18
|||||
Db 464 RCITPGSCRCCECNK 478

RESULT 25
BAD16739
ID BAD16739 PRELIMINARY; PRT; 2871 AA.
AC BAD16739;
DT 26-APR-2004 (TREMBlrel. 27, Created)
DT 26-APR-2004 (TREMBlrel. 27, Last sequence update)
DT 26-APR-2004 (TREMBlrel. 27, Last annotation update)
DE Fibrillin 1.
GN FN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueda T., Takahashi T., Eto S., Sato T., Xu G., Toki T., Yonesaka S.,
RA Ito E.;
```

DT 26-APR-2004 (TReMBLrel. 27, Last sequence update)
 DE 26-APR-2004 (TReMBLrel. 27, Last annotation update)
 FN Fibrillin 1.
 GN FBNI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Uyeda T., Takahashi T., Eto S., Sato T., Xu G., Toki T., Yonesaka S.,
 RA Ito E.;
 RT "3 Novel mutations of Fibrillin-1 and 10 single nucleotide
 RT polymorphisms of Fibrillin-3 in Marfan syndrome patients."
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; ABI77803; BAD16739.1; -
 SQ SEQUENCE 2871 AA; 312219 MW; C3D044A80321E082 CRC64;

Query Match 45.7%; Score 53; DB 2; Length 2871;
 Best Local Similarity 60.0%; Pred. No. 2e+02;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Qy 4 RCICITRGFCRCICTR 18
 ||| ||| ||| |||
 Db 464 RCITPGSCRCENK 478

RESULT 26
 Q7PJ70
 ID Q7PJ70 PRELIMINARY; PRT; 129 AA.
 AC Q7PJ70;
 DT 01-MAR-2004 (TReMBLrel. 26, Created)
 DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
 DE ENSANGP00000023822 (Fragment).
 GN Name=ENSANGP00000023822 (Fragment).
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
 OX NCBI_TaxID=180454;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PEST;
 RA Anopheles Genome Sequencing Consortium;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAAA01008964; EAA43873.1; -
 FT NON TER 1
 FT NON TER 129 129
 SQ SEQUENCE 129 AA; 13897 MW; C3D9F1EB9E65B74 CRC64;

Query Match 45.3%; Score 52.5; DB 2; Length 129;
 Best Local Similarity 50.0%; Pred. No. 16;
 Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
 Qy 3 RCICITRGFCRCICTR 18
 ||| ||| ||| |||
 Db 21 CSCVCVTRFL-CVCVR 35

RESULT 27
 Q94U26
 ID Q94U26 PRELIMINARY; PRT; 273 AA.
 AC Q94U26;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DE 28 kDa guide RNA binding protein.
 OS Leishmania tarentolae (Saurleishmania tarentolae).
 OG Mitochondrion.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania;

OC lizard Leishmania.
 OX NCBI_TaxID=5689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Aphasizhev R., Aphasizheva I., Simpson L.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF411047; AAL05568.1; -
 SQ SEQUENCE 273 AA; 31149 MW; 4C0321A819A9E564 CRC64;

Query Match 45.3%; Score 52.5; DB 2; Length 273;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 8; Conservative 2; Mismatches 3; Indels 3; Gaps 1;
 Qy 3 RCICITRGFCRCICTR 18
 ||| ||| ||| |||
 Db 253 CVCVCV--CVCVCCTR 265

RESULT 28
 Q6RY99
 ID Q6RY99 PRELIMINARY; PRT; 512 AA.
 AC Q6RY99;
 DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
 DE Gamma-hydroxybutyrate receptor.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Hippocampus;
 RA Andriampandry C., Taleb O., Viry S., Muller C., Humbert J.P.,
 RA Gobaille S., Aunis D., Maitre M.;
 RT "Cloning and characterization of a rat brain receptor that binds the
 RT endogenous neuromodulator gamma-hydroxybutyrate."
 RL FASEB J. 0:0-0(2004).
 DR EMBL; AY485933; AAR24072.1; -
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR001209; Ribosomal S14.
 DR InterPro; IPR008952; Tetraspanin.
 DR Pfam; PF00335; Tetraspanin; 1.
 DR PRINTS; PR00259; TMFOUR.
 DR PROSITE; PS00527; RIBOSOMAL_S14; UNKNOWN_1.
 KW Receptor.
 SQ SEQUENCE 512 AA; 56146 MW; 44DD12B937C4D060 CRC64;

Query Match 45.3%; Score 52.5; DB 2; Length 512;
 Best Local Similarity 50.0%; Pred. No. 52;
 Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
 Qy 3 RCICITRGFCRCICTR 18
 ||| ||| ||| |||
 Db 386 CVCVCVR-VCLCLCVR 400

RESULT 29
 AAR24072
 ID AAR24072 PRELIMINARY; PRT; 512 AA.
 AC AAR24072;
 DT 02-MAR-2004 (TReMBLrel. 27, Created)
 DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
 DE Gamma-hydroxybutyrate receptor.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Hippocampus;

RA Andriampandry C., Taleb O., Vary S., Muller C., Humbert J.P.,
RA Gobalet S., Anis D., Maître M.;
RA Cabañes J., Viret J., Viret J., Viret J., Viret J., Viret J.,
RT cell lines and characterization of a rat brain receptor that binds the
RT endogenous neuromodulator gamma-hydroxybutyrate.";
RL FASEB J. 0.9-0(2004).
RL FASEB J. 0.9-0(2004).
RL ENBL: AY485933; AAR24072.1; -.
RC RECEPTOR. 512 AA; 551146 MW; 44DD12B937C40D60 CRC64;
SC SOURCE.

| | | | | |
|-----------------------|--------------|--------------|---------------|-------------|
| Query Match | 45.3%; | Score 52.5; | DB 2; | Length 512; |
| Best Local Similarity | 50.8%; | Prod No. 52; | | |
| Matches 8; | Conservative | 2; | Mismatches 5; | Indels 1; |
| Gaps | 1; | | | |

| | | | |
|----|-----|-----------------|-----|
| Qy | 3 | CRGICRGFCRCICTR | 18 |
| | | | |
| Db | 386 | CVCVCVR-VCLCLVR | 400 |

```

RESULT 30
Q9PYQ3 PRELIMINARY; PRT; 66 AA.
ID Q9PYQ3
AC Q9PYQ3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE ORF143.
GN Name=ORF143;
OS Xestia c-nigrum granulosis virus (XnGV) (Xestia c-nigrum
  granulovirus).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX NCBI_TaxID=51677;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99434230; Pubmed=10502508;
RA Hayakawa T., Ko R., Okano K., Seong S.I., Goto C., Maeda S.;
RT "Sequence analysis of the Xestia c-nigrum granulovirus genome.";
RL Virology 262:277-297(1999).
DR EMBL; AF162221; AAF05257.1; --
SQ SEQUENCE 66 AA; 7457 MW; D814F5A3F0ABE10 CRC64;

```

```

Query Match      44.8%; Score 52; DB 2; Length 66;
Best Local Similarity 47.1%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GFCRCICTRGFCRCICT 17
Db 23 GCGCGRCCTKMCWCMT 39

```

```

RESULT 31
Q7QBV4
ID Q7QBV4 PRELIMINARY; PRT; 146 AA.
AC Q7QBV4;
DT 01-MAR-2004 (trEMBLrel. 26, Created)
DT 01-MAR-2004 (trEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (trEMBLrel. 26, Last annotation update)
DE AGCPI494 (Fragment).
GN Name=agCG47401; ORFNames=ENSANGG00000012913;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxId=180454;
[1]
RN PEST
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008859; EAA07639.1; -.
FT NON TER 1
SQ SEQUENCE 146 AA; 7D32B8AAABE776F98 CRC64;

```

```

Query Match      44.8%; Score 52; DB 2; Length 146;
Best Local Similarity 54.5%; Pred. No. 20;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      5  CICTRGFCRCI 15
      | | | | |
Db      44  CVCVLGWCRCV 54

```

```

RESULT 32
QBMZ55
QBMZ55 PRELIMINARY; PRT; 161 AA.
ID
AC Q8M255;
AC Q8M255;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GH14618P.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkely.
RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese E.,
RA George R., Gonzales M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Murgall C.J., Nunoo J., Facleb J., Paragas V., Park S.,
RA Patel S., Phouanenvong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Calniker S.;
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY113346; AAM29351.1; -.
DR FlyBase; FBgn0063249; BcDNA:GH14618.
SO SEQUENCE 161 AA; 17659 MW; EB41C9D94274EE94 CRC64;

```

```

Query Match      44.8%; Score 52; DB 2; Length 161;
Best Local Similarity 52.9%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1  GFCRCICRTGFCRCICCT 17
      ||||| : ||||| :
      ||||| : ||||| :
      115  GFCGSEIIISCVFCRCOCSS 131

db

```

```

RESULT 33
Q7PRP5 PRELIMINARY; PRT; 1823 AA.
AC Q7PRP5;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE ENSANGP00000019046 (Fragment).
GN Name=ENSANGG00000016557;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OC NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008847; EAA06854.2; -.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR001791; Laminin G.

```

```

DR InterPro; IPR001759; Pentaxin.
DR InterPro; IPR001211; PhospholipaseA2.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00008; EGF_5.
DR Pfam; PF02494; HIR; 2.
DR Pfam; PF00354; Pentaxin; 1.
DR Pfam; PF00084; Sushi; 8.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR00895; PENTAXIN.
DR PRINTS; PR00453; VWFADOMAIN.
DR PRODOM; PD002153; Pentaxin; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_5.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS00026; EGF_3; 5.
DR PROSITE; PS00825; HVR; 2.
DR PROSITE; PS00025; LAM_G DOMAIN; 1.
DR PROSITE; PS00118; PA2_HIS; UNKNOWN_1.
DR PROSITE; PS00923; SUSHI; 8.
DR PROSITE; PS00234; VWF_A; 1.
FT NON TER 1
FT TER 1823
SQ SEQUENCE 1823 AA; 200582 MW; 5740C4C700804379 CRC64;

Query Match 44.8%; Score 52; DB 2; Length 1823;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 CRCTCTRGFCRCIC 16
DB 263 CSLTTSYHYRCIC 276

RESULT 34
Q9UI23 PRELIMINARY; PRT; 190 AA.
AC Q9UI23;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE PRO0529.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE=Liver;
RA Zhang C., Yu Y., Zhang S., Ouyang S., Luo L., Wei H., Zhou G.,
RA Zhang Y., Liu M., He F.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111848; AAF16687.1; -
SQ SEQUENCE 190 AA; 21480 MW; 4B81042A29AA33844 CRC64;

Query Match 44.4%; Score 51.5; DB 2; Length 190;
Best Local Similarity 47.1%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 4; Indels 3; Gaps 1;

QY 3 CRCTCTRG---FCRCIC 16
DB 48 CLCMVRCGVSVVCVC 64

RESULT 35
Q6IG37 PRELIMINARY; PRT; 59 AA.
AC Q6IG37;
DT 03-JUL-2004 (TrEMBLrel. 27, Created)
DT 03-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 03-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HDC07368.
GN ORFNames=HDC07368;

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RC SEQUENCE FROM N.A.
RA Hild M., Beckmann B., Haas S., Koch B., Solov'yev V., Busold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Hoheisel J., Faro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome.";
RL Genome Biol. 5:R3-R3(2003).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK003929; DAA02627.1; -
DR InterPro; IPR001450; 4Fe4S ferredoxin.
DR PROSITE; PS00198; 4FE4S FERREDOXIN; UNKNOWN_1.
DR PROSITE; PS00198; 4FE4S FERREDOXIN; UNKNOWN_1.
SQ SEQUENCE 59 AA; 6339 MW; 7BD2C0B5C0F84905 CRC64;

Query Match 44.0%; Score 51; DB 2; Length 59;
Best Local Similarity 53.3%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CRCTCTRGFCRCIC 17
DB 43 CPCCQCQCECCCCCT 57

RESULT 36
IBB PHAAU STANDARD; PRT; 72 AA.
AC P01062;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Bowman-Birk type trypsin inhibitor.
OS Phaseolus aureus (Mung bean) (Vigna radiata).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eusoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
OX NCBI_TaxID=3916;
RN [1]
RC SEQUENCE.
RA MEDLINE=82249808; PubMed=6125033;
RA Zhang Y., Luo S., Tan F., Qi Z., Xu L., Zhang A.;
RT "Complete amino acid sequence of mung bean trypsin inhibitor.";
RL Sci. Sin., Ser. B, Chem. Biol. Agric. Med. Earth Sci. 25:268-277(1982).
RN [2]
RC X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 10-31, AND REVISIONS.
RA MEDLINE=95096014; PubMed=7798176;
RA Li Y., Huang Q., Zhang S., Liu S., Qi C., Tang Y.;
RT "Studies on an artificial trypsin inhibitor peptide derived from the
RT mung bean trypsin inhibitor: chemical synthesis, refolding, and
RT crystallographic analysis of its complex with trypsin.";
RL J. Biochem. 116:18-25(1994).
CC -!- MISCELLANEOUS: Functionally this inhibitor is unusual in that it
CC stoichiometrically inhibits trypsin in a molar ratio of 1:2.
CC -!- MISCELLANEOUS: The specificities and functions of this superfamily
CC of inhibitors depend not only on the active sites within the
CC domains, but also upon the amino acid composition, and resulting
CC molecular conformation, surrounding these regions.
CC -!- MISCELLANEOUS: Three isoforms are also found whose amino ends
CC differ slightly from that shown.
CC -!- SIMILARITY: Belongs to the Bowman-Birk serine protease inhibitor
CC family.
DR PIR; A01301; TIMB.
DR PDB; 1DF9; X-ray; C=1-72.
DR PDB; 1G91; X-ray; I=10-31.
DR PDB; 1SBW; X-ray; I=8-42.
DR PDB; 1SMF; X-ray; I=10-31.
DR InterPro; IPR000877; Prot. inh BBI.
DR Pfam; PF00228; Bowman-Birk_leg; 2.

```

DR ProDom; P0002168; Bowman-Birk_leg; 1.
DR SMART; SM00269; BowB; 1.
DR PROSITE; P500281; BOWMAN BIRK; 1.
KW 3D-structure; Direct protein sequencing; Serine protease inhibitor.
FT SITE 20 21
FT SITE 47 48
FT DISULFID 12 66
FT DISULFID 13 28
FT DISULFID 16 62
FT DISULFID 18 26
FT DISULFID 36 43
FT DISULFID 40 55
FT DISULFID 45 53
FT VARIANT 1 2
FT VARIANT 2 2
FT VARIANT 3 3
FT STRAND 17 17
FT STRAND 22 22
FT STRAND 26 27
FT STRAND 30 30
FT TURN 38 39
FT STRAND 43 45
FT STRAND 49 49
FT TURN 50 50
FT STRAND 53 55
FT STRAND 59 59
FT TURN 67 68
SQ SEQUENCE 72 AA; 7959 MW; 8359DE1A8E61E4F8 CRC64;

Query Match 44.0%; Score 51; DB 1; Length 72;
Best Local Similarity 57.9%; Pred. No. 15;

Matches 11; Conservative 2; Mismatches 2; Indels 4; Gaps 2;

Qy 3 CR-CICTR--GFCRCIC 17
Db 40 CKSCICTRSMFGKRCCLDT 58

RESULT 37

Q6ZQS2 ID Q6ZQS2 PRELIMINARY; PRT; 201 AA.
AC Q6ZQS2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ45585.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Thalamus;
RC Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128797; BAC87611.1; -.
DR InterPro; IPR06209; EGF-like.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
SQ SEQUENCE 201 AA; 20938 MW; 3408F8E817AA0500 CRC64;

Query Match 44.0%; Score 51; DB 2; Length 201;
Best Local Similarity 42.9%; Pred. No. 36;

Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 3 CR-CICTRGFCRCIC 16
Db 59 CLCVCLVSVCLVC 72

RESULT 38

BAC87611 ID BAC87611 PRELIMINARY; PRT; 201 AA.
AC BAC87611;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE CDNA FLJ45585 fis, clone BRTHA3013882.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Thalamus;
RC Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128797; BAC87611.1; -.
SQ SEQUENCE 201 AA; 20938 MW; 3408F8E817AA0500 CRC64;

Query Match 44.0%; Score 51; DB 2; Length 201;
Best Local Similarity 42.9%; Pred. No. 36;

Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 3 CR-CICTRGFCRCIC 16
Db 59 CLCVCLVSVCLVC 72

RESULT 39

Q91ZH6 ID Q91ZH6 PRELIMINARY; PRT; 326 AA.
AC Q91ZH6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Vascular endothelial growth factor C (Fragment).
GN Name=vegfc;
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC Meriones.
OC NCBI_TaxID=10047;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Lung;
RC Panilaitis B.J., Fuhrman J.A.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the PDGF/VEGF growth factor family.
DR EMBL; AF432867; AAL28127.1; -.
DR HSP; P01127; IPDG.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008083; F:growth factor activity; IEA.
DR GO; GO:0008151; P:cell growth and/or maintenance; IEA.
DR InterPro; IPR004153; CXKC repeat.
DR InterPro; IPR002400; GF_cycknot.
DR InterPro; IPR000072; PD_growth_factor.
DR Pfam; PF03128; CXKC; 5.
DR Pfam; PF00341; PDGF; 1.
DR PRINTS; PR00438; GFYSKNOT.
DR ProDom; P001629; PD growth_factor; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.

KW Growth factor; Mitogen.
 FT NON_TER 1
 SQ SEQUENCE 326 AA; 36826 MW; DOB27272C77836914 CRC64;
 Query Match 44.0%; Score 51; DB 2; Length 326;
 Best Local Similarity 36.0%; Pred. No. 55;
 Matches 9; Conservative 2; Mismatches 4; Indels 10; Gaps 1;
 Oy 3 CRCICTR-----GFCICICT 17
 Db 246 CQCVKRTCPRNQPLNFGKTCCT 270
 RESULT 40
 Q7PVN9
 ID Q7PVN9 PRELIMINARY; PRT; 336 AA.
 AC Q7PVN9;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE ENSANGP0000010520 (Fragment).
 GN Name=ENSANGS0000008031;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
 OX NCBI_TaxID=180454;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PEST;
 RA Anopheles Genome Sequencing Consortium;
 RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAB01008984; EAA14914.1; -.
 FT NON_TER 1
 FT NON_TER 336
 SQ SEQUENCE 336 AA; 36383 MW; 0C28D220B6E129A8 CRC64;
 Query Match 44.0%; Score 51; DB 2; Length 336;
 Best Local Similarity 55.6%; Pred. No. 56;
 Matches 10; Conservative 1; Mismatches 3; Indels 4; Gaps 2;
 Oy 3 CRCICTRGFC---RCICT 17
 Db 162 CGCTC-RGFCDFHRCACS 178

Search completed: October 26, 2004, 15:39:20
 Job time : 131.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 15:20:02 ; Search time 31.5 Seconds
(without alignments)
37.896 Million cell updates/sec

Title: US-10-009-317A-33

Perfect score: 112
Sequence: 1 GVCRLCRRGVCRCLCRR 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 56318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A-COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B-COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------|
| 1 | 99 | 88.4 | 18 | 3 | US-09-309-487-1 |
| 2 | 99 | 88.4 | 18 | 4 | US-09-967-808-1 |
| 3 | 95 | 84.8 | 18 | 3 | US-09-309-487-9 |
| 4 | 95 | 84.8 | 18 | 4 | US-09-967-808-9 |
| 5 | 92 | 82.1 | 18 | 4 | US-10-141-645-1 |
| 6 | 89 | 79.5 | 18 | 4 | US-10-141-645-2 |
| 7 | 89 | 79.5 | 18 | 4 | US-10-141-645-3 |
| 8 | 89 | 79.5 | 18 | 4 | US-10-141-645-4 |
| 9 | 88 | 78.6 | 18 | 4 | US-10-141-645-5 |
| 10 | 88 | 78.6 | 18 | 4 | US-10-141-645-6 |
| 11 | 85 | 75.9 | 18 | 4 | US-10-141-645-7 |
| 12 | 80 | 71.4 | 18 | 4 | US-09-917-340-53 |
| 13 | 77 | 68.8 | 18 | 4 | US-10-141-645-8 |
| 14 | 77 | 68.8 | 18 | 4 | US-10-141-645-9 |
| 15 | 62 | 55.4 | 92 | 3 | US-09-309-487-22 |
| 16 | 62 | 55.4 | 92 | 4 | US-09-967-808-22 |
| 17 | 62 | 55.4 | 141 | 4 | US-10-141-645-68 |
| 18 | 62 | 55.4 | 141 | 4 | US-10-141-645-72 |
| 19 | 61 | 54.5 | 140 | 4 | US-10-141-645-66 |
| 20 | 61 | 54.5 | 141 | 4 | US-10-141-645-67 |
| 21 | 60 | 53.6 | 76 | 4 | US-09-309-487-16 |
| 22 | 60 | 53.6 | 76 | 4 | US-09-967-808-16 |
| 23 | 60 | 53.6 | 76 | 4 | US-10-141-645-17 |
| 24 | 60 | 53.6 | 92 | 3 | US-09-309-487-21 |
| 25 | 60 | 53.6 | 92 | 4 | US-09-967-808-21 |
| 26 | 60 | 53.6 | 141 | 4 | US-10-141-645-73 |
| 27 | 58 | 51.8 | 17 | 3 | US-09-604-864-1 |

Sequence 1, Appli
Sequence 266, App
Sequence 20, Appl
Sequence 20, Appl
Sequence 27, Appl
Sequence 45, Appl
Sequence 121, App
Sequence 65, Appl
Sequence 89, Appl
Sequence 125, App
Sequence 40304, A
Sequence 5520, A
Sequence 33, Appl
Sequence 12, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-309-487-1

; Sequence 1, Application US/09309487
; Patent No. 6335318

; GENERAL INFORMATION:

; APPLICANT: Selssted, Michael E.

; APPLICANT: Tang, Yi-Quan

; APPLICANT: Yuan, Jun

; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same

; FILE REFERENCE: P-UC 3095

; CURRENT APPLICATION NUMBER: US/09/309,487

; CURRENT FILING DATE: 1999-05-10

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Macaca mulatta

; US-09-309-487-1

Query Match 88.4%; Score 99; DB 3; Length 18;

Best Local Similarity 83.3%; Pred. No. 8.6e-06;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GVCRLCRRGVCRCLCRR 18

Db 1 GVCRLCRRGVCRCLCRR 18

RESULT 2

US-09-967-808-1

; Sequence 1, Application US/09967808

; Patent No. 6514727

; GENERAL INFORMATION:

; APPLICANT: Selssted, Michael E.

; APPLICANT: Tang, Yi-Quan

; APPLICANT: Yuan, Jun

; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using

; FILE REFERENCE: P-UC 3095

; CURRENT APPLICATION NUMBER: US/09/967,808

; CURRENT FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: US/09/309,487

; PRIOR FILING DATE: 1999-05-10

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 18

```
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-967-808-1

Query Match      88.4%; Score 99; DB 4; Length 18;
Best Local Similarity 83.3%; Pred. No. 8.6e-06;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVCRLCRRGVCRCLCR 18
Db 1 GFCRLCRRGVCRCLCTR 18

RESULT 3
US-09-309-487-9
; Sequence 9, Application US/09309487
; Patent No. 6335318
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Quellet, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/309,487
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-309-487-9

Query Match      84.8%; Score 95; DB 3; Length 18;
Best Local Similarity 87.5%; Pred. No. 2.5e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVCRLCRRGVCRCLC 16
Db 3 GFCRLCRRGVCRCLC 18

RESULT 4
US-09-967-808-9
; Sequence 9, Application US/09967808
; Patent No. 6514727
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Quellet, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/967,808
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-967-808-9

Query Match      84.8%; Score 95; DB 4; Length 18;
Best Local Similarity 87.5%; Pred. No. 2.5e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVCRLCRRGVCRCLC 16
Db 1 GFCRLCRRGVCRCLCTR 18

; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-967-808-1

Query Match      88.4%; Score 99; DB 4; Length 18;
Best Local Similarity 83.3%; Pred. No. 8.6e-06;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVCRLCRRGVCRCLCR 18
Db 1 GFCRLCRRGVCRCLCTR 18

RESULT 5
US-10-141-645-1
; Sequence 1, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-141-645-1

Query Match      82.1%; Score 92; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 5.6e-05;
Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVCRLCRRGVCRCLCR 18
Db 1 GICRCICRGICRCICGR 18

RESULT 6
US-10-141-645-2
; Sequence 2, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic variant
US-10-141-645-2

Query Match      79.5%; Score 89; DB 4; Length 18;
Best Local Similarity 61.1%; Pred. No. 0.00012;
Matches 11; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVCRLCRRGVCRCLCR 18
Db 1 GICRCICRGICRCICGR 18
```


RESULT 19
US-10-141-645-66
; Sequence 66, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:

; APPLICANT: Robert Lehrer
 ; APPLICANT: Alan Waring
 ; APPLICANT: Alexander Cole
 ; APPLICANT: Teresa Hong
 ; TITLE OF INVENTION: Retrocyclins - Antiviral and
 ; FILE REFERENCE: UCLA-001CIP
 ; CURRENT APPLICATION NUMBER: US/10/141,645
 ; CURRENT FILING DATE: 2002-05-06
 ; PRIOR APPLICATION NUMBER: 60/284,855
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: Unassigned
 ; PRIOR FILING DATE: 2002-04-18
 ; NUMBER OF SEQ ID NOS: 125
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 66
 ; LENGTH: 140
 ; TYPE: PRT
 ; ORGANISM: Orangutan
 US-10-141-645-66

Query Match 54.5%; Score 61; DB 4; Length 140;
 Best Local Similarity 83.3%; Pred. No. 1.2;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCLCRGVCRCRL 15
 ||:|||||||
 Db 68 RCLCRGVCRCFL 79

RESULT 20
 US-10-141-645-67
 ; Sequence 67, Application US/10141645
 ; Patent No. 6713078
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert Lehrer
 ; APPLICANT: Alan Waring
 ; APPLICANT: Alexander Cole
 ; APPLICANT: Teresa Hong
 ; TITLE OF INVENTION: Retrocyclins - Antiviral and
 ; FILE REFERENCE: UCLA-001CIP
 ; CURRENT APPLICATION NUMBER: US/10/141,645
 ; CURRENT FILING DATE: 2002-05-06
 ; PRIOR APPLICATION NUMBER: 60/284,855
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: Unassigned
 ; PRIOR FILING DATE: 2002-04-18
 ; NUMBER OF SEQ ID NOS: 125
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 67
 ; LENGTH: 141
 ; TYPE: PRT
 ; ORGANISM: Orangutan
 US-10-141-645-67

Query Match 54.5%; Score 61; DB 4; Length 141;
 Best Local Similarity 83.3%; Pred. No. 1.2;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCLCRGVCRCRL 15
 ||:|||||||
 Db 69 RCLCRGVCRCFL 80

RESULT 21
 US-09-309-487-16
 ; Sequence 16, Application US/09309487
 ; Patent No. 6335318
 ; GENERAL INFORMATION:
 ; APPLICANT: Selsted, Michael E.
 ; APPLICANT: Tang, Yi-Quan
 ; APPLICANT: Yuan, Jun

; APPLICANT: Ouellette, Andre J.
 ; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
 ; FILE REFERENCE: P-UC 3095
 ; CURRENT APPLICATION NUMBER: US/09/309,487
 ; CURRENT FILING DATE: 1999-05-10
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 16
 ; LENGTH: 76
 ; TYPE: PRT
 ; ORGANISM: Macaca mulatta
 US-09-309-487-16

Query Match 53.6%; Score 60; DB 3; Length 76;
 Best Local Similarity 83.3%; Pred. No. 0.94;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCLCRGVCRCRL 15
 ||:|||||||
 Db 65 RCLCRGVCQLL 76

RESULT 22
 US-09-967-808-16
 ; Sequence 16, Application US/09967808
 ; Patent No. 6514727
 ; GENERAL INFORMATION:
 ; APPLICANT: Selsted, Michael E.
 ; APPLICANT: Tang, Yi-Quan
 ; APPLICANT: Yuan, Jun
 ; APPLICANT: Ouellette, Andre J.
 ; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
 ; FILE REFERENCE: P-UC 3095
 ; CURRENT APPLICATION NUMBER: US/09/967,808
 ; CURRENT FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: US/09/309,487
 ; PRIOR FILING DATE: 1999-05-10
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 16
 ; LENGTH: 76
 ; TYPE: PRT
 ; ORGANISM: Macaca mulatta
 US-09-967-808-16

Query Match 53.6%; Score 60; DB 4; Length 76;
 Best Local Similarity 83.3%; Pred. No. 0.94;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCLCRGVCRCRL 15
 ||:|||||||
 Db 65 RCLCRGVCQLL 76

RESULT 23
 US-10-141-645-17
 ; Sequence 17, Application US/10141645
 ; Patent No. 6713078
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert Lehrer
 ; APPLICANT: Alan Waring
 ; APPLICANT: Alexander Cole
 ; APPLICANT: Teresa Hong
 ; TITLE OF INVENTION: Retrocyclins - Antiviral and
 ; FILE REFERENCE: UCLA-001CIP
 ; CURRENT APPLICATION NUMBER: US/10/141,645
 ; CURRENT FILING DATE: 2002-05-06
 ; PRIOR APPLICATION NUMBER: 60/284,855
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: Unassigned
 ; PRIOR FILING DATE: 2002-04-18

; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Macaca mulatta
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
US-10-141-645-17

Query Match 53.6%; Score 60; DB 4; Length 76;
Best Local Similarity 83.3%; Pred. No. 0.94; 1; Indels 0; Gaps 0;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RCLCRGVCRCL 15
|||:|||||:
Db 65 RCICRGVCQLL 76

RESULT 24

US-09-309-487-21
; Sequence 21, Application US/09309487
; Patent No. 6335318
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/309,487
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-309-487-21

Query Match 53.6%; Score 60; DB 3; Length 92;
Best Local Similarity 69.2%; Pred. No. 1.1; 3; Indels 0; Gaps 0;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 RCLCRGVCRCL 16
|||:|||||:
Db 65 RCICRGVCRLC 77

RESULT 25

US-09-967-808-21
; Sequence 21, Application US/09967808
; Patent No. 6514727
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/967,808
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-967-808-21

Query Match 53.6%; Score 60; DB 4; Length 92;
Best Local Similarity 69.2%; Pred. No. 1.1; 3; Indels 0; Gaps 0;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 RCLCRGVCRCL 16
|||:|||||:
Db 65 RCICRGVCRLC 77

RESULT 26

US-10-141-645-73
; Sequence 73, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; TITLE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-COICIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Pig-tailed macaque
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 113
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-141-645-73

Query Match 53.6%; Score 60; DB 4; Length 141;
Best Local Similarity 71.4%; Pred. No. 1.6; 2; Indels 0; Gaps 0;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 RCLCRGVCRCL 17
|||:|||||:
Db 69 RCICRGVCQLLR 82

RESULT 27

US-09-604-864-1
; Sequence 1, Application US/09604864
; Patent No. 6337317
; GENERAL INFORMATION:
; APPLICANT: Hancock, Robert E. W.
; APPLICANT: Zhang, Lijuan
; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: UBC1170
; CURRENT APPLICATION NUMBER: US/09/604,864
; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Limulus polyphemus
US-09-604-864-1

Query Match 51.8%; Score 58; DB 3; Length 17;
Best Local Similarity 62.5%; Pred. No. 0.47; 5; Indels 0; Gaps 0;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CRCLCRGVCRCLCR 18
Db 2 CFAVCRGRCRYKCR 17

RESULT 28

US-10-042-872-1
; Sequence 1, Application US/10042872
; Patent No. 6747007
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF BRITISH COLUMBIA
; APPLICANT: Hancock, Robert E. W.
; APPLICANT: Zhang, Lijuan
; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES AND METHODS OF
; FILE REFERENCE: USE THEREOF
; CURRENT APPLICATION NUMBER: US/10/042,872
; CURRENT FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 09/604,864
; PRIOR FILING DATE: 2000-06-27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Limulus polyphemus
US-10-042-872-1

Query Match 51.8%; Score 58; DB 4; Length 17;
Best Local Similarity 62.5%; Pred. No. 0.47;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CRCLCRGVCRCLCR 18
Db 2 CFAVCRGRCRYKCR 17

RESULT 29

US-09-510-238A-286
; Sequence 286, Application US/09510238A
; Patent No. 6730303
; GENERAL INFORMATION:
; APPLICANT: FENG, YIQUING
; APPLICANT: BAUM, CHARLES M
; APPLICANT: CAPARON, MAIRE H
; APPLICANT: ZURFLUH, LINDA L
; APPLICANT: KLEIN, BARBARA K
; APPLICANT: MCWHERTER, CHARLES A
; APPLICANT: STATEN, NICHOLAS R
; APPLICANT: SUMMERS, NEENA L
; APPLICANT: BAUER, S C
; APPLICANT: LEE, STEPHEN C
; TITLE OF INVENTION: MULTI-FUNCTIONAL HEMATOPOIETIC RECEPTOR
; AGONISTS
; NUMBER OF SEQUENCES: 313
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DENNIS A. BENNETT, G.D. SEARLE & CO.,
; CORPORATE PATENT DEPT.
; STREET: P.O. BOX 5110
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/510,238A
; FILING DATE: 22-Feb-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/US 96/15774
; FILING DATE: 06-OCT-1996
; APPLICATION NUMBER: US 60/004,834
; FILING DATE: 05-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BENNETT, DENNIS A
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: 2910/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-694-5402
; TELEFAX: 314-694-9095
; INFORMATION FOR SEQ ID NO: 286:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 286:
US-09-510-238A-286

Query Match 50.9%; Score 57; DB 4; Length 180;
Best Local Similarity 50.0%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GVCRLCRGVCRCLC 16
Db 8 GACCCACTTGCCTC 23

RESULT 30

US-09-309-487-20
; Sequence 20, Application US/09309487
; Patent No. 6335318
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/309,487
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 20
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-309-487-20

Query Match 50.0%; Score 56; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RCLCRGVGC 12
Db 1 RCLCRGVGC 9

RESULT 31

US-09-967-808-20
; Sequence 20, Application US/09967808
; Patent No. 6514727
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
; FILE REFERENCE: P-UC 3095

; CURRENT APPLICATION NUMBER: US/09/967,808
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-967-808-20

Query Match 50.0%; Score 56; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RCLCRGVC 12
Db 1 RCLCRGVC 9

RESULT 32
US-10-141-645-27
; Sequence 27, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; PRIOR FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: generated by replacement of variants in consensus
; OTHER INFORMATION: sequence
US-10-141-645-27

Query Match 50.0%; Score 56; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RCLCRGVC 12
Db 1 RCLCRGVC 9

RESULT 33
US-10-141-645-45
; Sequence 45, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06

; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: generated by replacement of variants in consensus
; OTHER INFORMATION: sequence
US-10-141-645-45

Query Match 50.0%; Score 56; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RCLCRGVC 12
Db 1 RCLCRGVC 9

RESULT 34
US-10-141-645-121
; Sequence 121, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-141-645-121

Query Match 50.0%; Score 56; DB 4; Length 118;
Best Local Similarity 64.3%; Pred. No. 3.9;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RCLCRGVCRLCR 17
Db 46 RCICRGICRLRR 59

RESULT 35
US-10-141-645-65
; Sequence 65, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06

; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-141-645-65

Query Match 50.0%; Score 56; DB 4; Length 140;
Best Local Similarity 64.3%; Pred. No. 4.5;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 RCLCRGGVCRCLCR 17
||:|||||
DB 68 RCICRGICRLRR 81

RESULT 36
US-10-141-645-69
; Sequence 69, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Gorilla
US-10-141-645-69

Query Match 50.0%; Score 56; DB 4; Length 140;
Best Local Similarity 64.3%; Pred. No. 4.5;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 RCLCRGGVCRCLCR 17
||:|||||
DB 68 RCICRGICRLRR 81

RESULT 37
US-10-141-645-125
; Sequence 125, Application US/10141645
; Patent No. 6713078
; GENERAL INFORMATION:
; APPLICANT: Robert Lehrer
; APPLICANT: Alan Waring
; APPLICANT: Alexander Cole
; APPLICANT: Teresa Hong
; TITLE OF INVENTION: Retrocyclins - Antiviral and
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned

; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 125
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-141-645-125

Query Match 50.0%; Score 56; DB 4; Length 141;
Best Local Similarity 64.3%; Pred. No. 4.5;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 RCLCRGGVCRCLCR 17
||:|||||
DB 69 RCICRGICRLRR 82

RESULT 38
US-09-252-991A-32502
; Sequence 32502, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32502
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32502

Query Match 49.1%; Score 55; DB 4; Length 168;
Best Local Similarity 66.7%; Pred. No. 6.8;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 CLCRGGVCRCLC 16
||:|||||
DB 150 CRCRGPCRCAC 161

RESULT 39
US-09-270-767-40304
; Sequence 40304, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40304
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-40304

Query Match 48.7%; Score 54.5; DB 4; Length 96;
Best Local Similarity 47.4%; Pred. No. 4.9;
Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 3 CRCLCR---GVCRCLCR 18
||:|||||

Db 16 CGCVCORRCAGVCAECCK 34

RESULT 40
US-09-270-767-55520
; Sequence 55520, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55520
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-55520

Query Match 48.7%; Score 54.5; DB 4; Length 96;
Best Local Similarity 47.4%; Pred. No. 4.9;
Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 3 CRCLCRR--GVCRLCRR 18
| | | | | | | | | | | | | | | | | | | | | |
Db 16 CGCVCORRCAGVCAECCK 34

Search completed: October 26, 2004, 15:41:19
Job time : 32.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 15:15:31 ; Search time 23.5 Seconds

(without alignments)
73.698 Million cell updates/sec

Title: US-10-009-317A-33

Perfect score: 112

Sequence: 1 GVCRLCRGVCRCLCRR 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 80 | 71.4 | 18 | 2 C59089 | theta defensin-1 - |
| 2 | 50 | 53.6 | 76 | 2 B59089 | theta defensin lb |
| 3 | 58 | 51.8 | 188 | 2 T15651 | hypothetical prote |
| 4 | 57 | 50.9 | 164 | 2 T24272 | hypothetical prote |
| 5 | 53.5 | 47.8 | 419 | 2 S69207 | vascular endotheli |
| 6 | 52.5 | 46.9 | 1131 | 2 T15617 | hypothetical prote |
| 7 | 52 | 46.4 | 152 | 2 T18975 | hypothetical prote |
| 8 | 51 | 45.5 | 76 | 2 A59089 | theta defensin la |
| 9 | 51 | 45.5 | 85 | 2 T48125 | hypothetical prote |
| 10 | 50.5 | 45.1 | 582 | 2 B45878 | hypothetical prote |
| 11 | 50 | 44.6 | 131 | 2 S50807 | probable membrane |
| 12 | 49.5 | 44.2 | 248 | 2 E71602 | probable integral |
| 13 | 49 | 43.8 | 476 | 2 JC5042 | G protein-coupled |
| 14 | 48 | 42.9 | 290 | 2 J72858 | AcOrf-70 protein - |
| 15 | 48 | 42.9 | 303 | 2 B70554 | hypothetical prote |
| 16 | 48 | 42.9 | 1217 | 1 EGM5MG | epidermal growth f |
| 17 | 47.5 | 42.4 | 265 | 2 A41116 | transcription fact |
| 18 | 47 | 42.0 | 1700 | 2 S08167 | Balbani ring 3 pr |
| 19 | 47 | 42.0 | 2206 | 1 GNNY21 | genome polyprotein |
| 20 | 47 | 42.0 | 2910 | 2 T42214 | otogelin - mouse |
| 21 | 46.5 | 41.5 | 77 | 2 I48725 | Q300 protein - mou |
| 22 | 46.5 | 41.5 | 256 | 2 T46871 | C-8 sterol isomera |
| 23 | 46.5 | 41.5 | 835 | 2 JP0076 | nel protein - chic |
| 24 | 46 | 41.1 | 53 | 2 S29214 | neurotoxin Tx2 - s |
| 25 | 46 | 41.1 | 79 | 1 LNP01 | pulmonary surfacta |
| 26 | 46 | 41.1 | 94 | 2 C37057 | fibronectin recept |
| 27 | 46 | 41.1 | 128 | 2 D72481 | hypothetical prote |
| 28 | 46 | 41.1 | 302 | 2 A25854 | chloramphenicol re |
| 29 | 46 | 41.1 | 321 | 2 H90942 | probable diogenase |

ALIGNMENTS

RESULT 1

CS9089

theta defensin-1 - rhesus macaque

N:Alternate names: RTD-1

C:Species: Macaca mulatta (rhesus macaque)

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: C59089

R:Tang, Y.Q.; Yuan, J.; Osapay, G.; Osapay, K.; Tran, D.; Miller, C.J.; Ouellette, A.J.;

Science 286, 498-502, 1999

A:Title: A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of F1-9/10/Cross-link: cyclopeptide (Cys-Arg) #status experimental

A:Reference number: A59089; MUID:99453140; PMID:10521339

A:Accession: C59089

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-18 <SEL>

A:Note: This sequence is cyclically permuted by -6 residues from the sequence presented i

C:Comment: For the two contributing precursor sequences, see PIR:A59089 and PIR:B59089.

C:Keywords: antibacterial; antibiotic; antifungal; leukocyte; protein splicing

F1-9/Region: theta defensin la-derived

F10-18/Region: theta defensin lb-derived

F11-18/Cross-link: cyclopeptide (Arg-Cys) #status experimental

F12-11,4-9,13-18/Disulfide bonds: #status experimental

F19-10/Cross-link: cyclopeptide (Cys-Arg) #status experimental

Query Match 71.4%; Score 80; DB 2; Length 18;

Best Local Similarity 80.0%; Pred. No. 0.0015;

Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RCLCRGVCRCLCRR 18

Db 1 RCICTRGFCRLCRR 15

RESULT 2

BS9089

theta defensin lb precursor - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: B59089

R:Tang, Y.Q.; Yuan, J.; Osapay, G.; Osapay, K.; Tran, D.; Miller, C.J.; Ouellette, A.J.;

Science 286, 498-502, 1999

A:Title: A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of

A:Reference number: A59089; MUID:99453140; PMID:10521339

A:Accession: B59089

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-76 <TAM>

A:Cross-references: UNIPROT:P82271; GB:AF191101; NID:96137229; PIDN:AAF04390.1; PID:96137

C:Comment: For the complete mature sequence, see PIR:C59089.

C:Superfamily: mammalian defensin

C:Keywords: antibacterial; antibiotic; antifungal; leukocyte; protein splicing

F1-20/Domain: signal sequence #status predicted <SIG>

F;21-64/Domain: amino-terminal propeptide #status predicted <PRO>
F;74-76/Domain: carboxyl-terminal propeptide #status predicted <CTP>

Query Match 53.6%; Score 60; DB 2; Length 76;
Best Local Similarity 83.3%; Pred. No. 0.81;
Matches 10; Conservative 1; Mismatches 0; Gaps 0;

Qy 4 RCLCRGVCRCL 15
| | | | | | | | | |
Db 65 RCLCRGVQCQLL 76

RESULT 3

Ti1561
hypothetical protein C27A2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T15651

R;Nhan, M.

submitted to the EMBL Data Library, May 1996

A;Description: The sequence of C. elegans cosmid C27A2.

A;Reference number: Z18382

A;Accession: T15651

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-188 <NHA>

A;Cross-references: UNIPROT:Q18238; EMBL:U58760; NID:g1330384; PID:g1330389; PIDN:AAB007

A;Experimental source: strain Bristol N2; clone C27A2

C;Genetics:

A;Gene: CESP:C27A2.5

A;Map position: 2

A;Introns: 19/3; 91/2

Query Match 51.8%; Score 58; DB 2; Length 188;
Best Local Similarity 55.6%; Pred. No. 2.4;
Matches 10; Conservative 8; Indels 0; Gaps 0;

Qy 1 GVCRCCLCRGVCRCLCR 18
| | | | | | | | | |
Db 81 GGCGCCCCRPKCCCCRR 98

RESULT 4

T24272
hypothetical protein T01B7.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24272

R;Sims, M.

submitted to the EMBL Data Library, October 1995

A;Reference number: Z19867

A;Accession: T24272

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-164 <WIL>

A;Cross-references: UNIPROT:Q22048; EMBL:Z66499; PIDN:CAA91301.1; GSPDB:GN00020; CESP:TO

A;Experimental source: clone T01B7

C;Genetics:

A;Gene: CESP:T01B7.8

A;Map position: 2

A;Introns: 20/3; 90/2

Query Match 50.9%; Score 57; DB 2; Length 164;
Best Local Similarity 55.6%; Pred. No. 2.9;
Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GVCRCCLCRGVCRCLCR 18
| | | | | | | | | |
Db 80 GGCGCCCCRRPRCCCCRR 97

RESULT 5

S69207

vascular endothelial growth factor C precursor - human
N;Alternate names: FLT4 ligand DHM

C;Species: Homo sapiens (man)

C;Date: 27-Apr-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004

C;Accession: S69207; S61795; S71443; S69208; G02659

R;Joukov, V.; Pajusola, K.; Kaipainen, A.; Chilov, D.; Lahtinen, I.; Kukkk, E.; Saksela, C.

EMBL J. 15, 1751, 1996

A;Title: Corrigendum: A novel vascular endothelial growth factor, VEGF-C, is a ligand for

A;Reference number: S69207; MUID:96203094; PMID:8612600

A;Accession: S69207

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-419 <JOU>

A;Cross-references: UNIPROT:P49767; EMBL:X94216; NID:g1177488; PIDN:CAA63907.1; PID:e2210

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995

A;Note: only a part of the translation is shown

A;Note: this is a revision to the sequence from reference S61795

R;Joukov, V.; Pajusola, K.; Kaipainen, A.; Chilov, D.; Lahtinen, I.; Kukkk, E.; Saksela, C.

EMBL J. 15, 290-298, 1996

A;Title: A novel vascular endothelial growth factor, VEGF-C, is a ligand for the Flt-4 (VE

A;Reference number: S61795; MUID:96178224; PMID:8617204

A;Accession: S61795

A;Status: nucleic acid sequence not shown; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 70-419 <JOU>

A;Note: this sequence has been revised in reference S69207

A;Accession: S71443

A;Molecule type: protein

A;Residues: 'X', 104-120 <JOU>

R;Lee, J.; Gray, A.; Yuan, J.; Luch, S.M.; Avraham, H.; Wood, W.I.

submitted to the EMBL Data Library, December 1995

A;Description: Vascular endothelial growth factor related protein (VRP): A ligand and spe

A;Reference number: S69208

A;Accession: S69208

A;Molecule type: mRNA

A;Residues: 1-419 <LSE>

A;Cross-references: EMBL:U43142; NID:g1150988; PIDN:AAA85214.1; PID:g1150989

R;Morris, J.C.

submitted to the EMBL Data Library, May 1996

A;Reference number: H01557

A;Accession: G02659

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-419 <MOR>

A;Cross-references: EMBL:U58111; NID:g1373426; PIDN:AA02909.1; PID:g1373427

C;Genetics:

A;Gene: GDB:VEGFC; VRP

A;Cross-references: GDB:3890883; OMIM:601528

F;1-12/Domain: signal sequence #status predicted <SIG>

F;13-102/Domain: propeptide #status predicted <PRO>

F;103-419/Product: vascular endothelial growth factor C #status experimental <MAT>

Query Match 47.8%; Score 53.5; DB 2; Length 419;

Best Local Similarity 26.7%; Pred. No. 14;

Matches 8; Conservative 6; Mismatches 1; Indels 15; Gaps 1;

Qy 3 CQCVCRGV-----CRCLCR 17

| | | | | | | | | |

Db 291 CQCVCRAGLRPASCGPHKELDRNSCQCVCK 320

RESULT 6

T15617

hypothetical protein C25F6.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C;Accession: T15617

R;Bentley, D.

submitted to the EMBL Data Library, October 1995

A;Description: The sequence of C. elegans cosmid C25F6.

A;Reference number: Z18377

A;Accession: T15617

A;Status: preliminary; translated from GB/EMBL/DBJ

```

A:Molecule type: DNA
A:Residues: 1-1131 <BEN>
A:Cross-references: EMBL:U39742; NID:G1049455; PID:G1049459; PIDN:AAA80434.1; CESP:C25F6
C:Genetics:
A:Gene: CESP:C25F6.2
A:Introns: 105/1; 210/3; 283/3; 316/1; 346/2; 463/3; 566/1; 722/2; 897/1; 991/2; 1032/1;
Query Match 46.9%; Score 52.5; DB 2; Length 1131;
Best Local Similarity 47.1%; Pred. No. 33;
Matches 8; Conservative 4; Mismatches 2; Indels 3; Gaps 1;
QY 2 VCRCLRRGVCRCLCR 18
| | | | | | | | | |
Db 5 LCVVCV---VCACVCRK 18

RESULT 7
T18975
Hypothetical protein C06A1.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18975
R:McMurray, A.
Submitted to the EMBL Data Library, June 1995
A:Reference number: Z19054
A:Accession: T18975
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-152 <WIL>
A:Cross-references: UNIPROT:Q9XVX3; EMBL:Z49886; PIDN:CAA90055.1; GSPDB:GN00020; CESP:C0
A:Experimental source: clone C06A1
C:Genetics:
A:Gene: CESP:C06A1.6
A:Map position: 2
A:Introns: 22/3

Query Match 46.4%; Score 52; DB 2; Length 152;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 GVCRCRCRGVCRCLCR 17
| | | | | | | | | |
Db 67 GGGCCCCCPRCCCCCR 83

RESULT 8
A59089
theta defensin la precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: A59089
R:Tang, Y.Q.; Yuan, J.; Osapay, G.; Tran, D.; Miller, C.J.; Ouellette, A.J.;
Science 286, 498-502, 1999
A:Title: A cyclic antimicrobial peptide produced in primate leukocytes by the ligation o
A:Reference number: A59089; MUID:99453140; PMID:10521339
A:Accession: A59089
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-76 <TAN>
A:Cross-references: UNIPROT:P82270; GB:AF191100; NID:G6137227; PIDN:AAF04389.1; PID:G613
C:Comment: For the complete mature sequence, see PIR:C59089.
C:Superfamily: mammalian defensin
C:Keywords: antibacterial; antibiotic; antifungal; leukocyte; protein splicing
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-64/Domain: amino-terminal propeptide #status predicted <PRO>
F;74-76/Domain: carboxyl-terminal propeptide #status predicted <CTP>

Query Match 45.5%; Score 51; DB 2; Length 76;
Best Local Similarity 66.7%; Pred. No. 8;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 4 RCLCRRGVCRCL 15
| | | | | | | |

```

```

Db 65 RCICTRGFCRL 76

RESULT 9
T48125
Hypothetical protein F16M2.210 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T48125
R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.
Submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24459
A:Accession: T48125
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-85 <RIE>
A:Cross-references: UNIPROT:Q9M1V4; EMBL:AL138648
A:Experimental source: cultivar Columbia; BAC clone F16M2
C:Genetics:
A:Map position: 3
A:Introns: 20/1
A:Note: F16M2.210

Query Match 45.5%; Score 51; DB 2; Length 85;
Best Local Similarity 44.4%; Pred. No. 9.6;
Matches 12; Conservative 2; Mismatches 3; Indels 10; Gaps 2;
QY 1 GVCRCRCR-----RGVC-----RCLCR 17
| | | | | | | | | |
Db 52 GLCDNLCKYEGALSGVCVSPHRCRCR 78

RESULT 10
B45878
Hypothetical protein 2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: B45878
R:Sarvetnick, N.; Tsai, J.Y.; Fox, H.; Pilder, S.H.; Silver, L.M.
Immunogenetics 30, 34-41, 1989
A:Title: A mouse chromosome 17 gene encodes a testes-specific transcript with unusual pr
A:Reference number: A45878; MUID:89307395; PMID:2568335
A:Accession: B45878
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-582 <SAR>
A:Cross-references: UNIPROT:Q62295; GB:M28921
A:Note: this sequence has been corrected in Immunogenetics 31, 283-284 (1990)
C:Superfamily: mouse hypothetical protein 2

Query Match 45.1%; Score 50.5; DB 2; Length 582;
Best Local Similarity 56.2%; Pred. No. 37;
Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
QY 2 VCRCLRRGVCRCLCR 17
| | | | | | | | | |
Db 71 VCSMCVY-VVCVCVR 85

RESULT 11
S50807
probable membrane protein YJL064w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein HRC131; hypothetical protein J1120
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1995 #sequence_revision 08-Sep-1995 #text_change 09-Jul-2004
C:Accession: S50807; S47126; S56938
R:Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
Yeast 11, 57-60, 1995
A:Title: Sequence of a 17.1 kb DNA fragment from chromosome X of Saccharomyces cerevisiae
A:Reference number: S50798; MUID:95282514; PMID:7762302
A:Accession: S50807
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

```

A;Residues: 1-131 <VAN>
A;Cross-references: UNIPROT:P47036; EMBL:Z34288; NID:g498992; PIDN:CAA84058.1; PID:g4990
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
R;Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
A;Submitted to the EMBL Data Library, June 1994
A;Description: Sequence analysis of a 17.1 kb DNA fragment from chromosome X of Saccharo
A;Reference number: S47117
A;Accession: S47126
A;Molecule type: DNA
A;Residues: 1-131 <VAV>
A;Cross-references: EMBL:Z34288; NID:g498992; PID:g499002
R;Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
A;Submitted to the Protein Sequence Database, September 1995
A;Reference number: S56835
A;Accession: S56838
A;Molecule type: DNA
A;Residues: 1-131 <POH>
A;Cross-references: EMBL:Z49340; NID:g1008212; PID:g1008214; MIPS:YJL064w
C;Genetics:
A;Map position: 10L
A;Superfamily: Saccharonocyes probable membrane protein YJL064w
C;Keywords: transmembrane protein

Query Match 44.6%; Score 50; DB 2; Length 131;
Best Local Similarity 52.9%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

Qy 1 GVCRLCRGGVCRCLCR 17
Db 55 GTCCGCC---CCCLCR 67

RESULT 12
E71502
Probable integral membrane protein PFB0950w - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C;Accession: E71502
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
P;Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A;Reference number: A71600; MUID:99021743; PMID:9804551
A;Accession: E71602
A;Status: preliminary;
A;Molecule type: DNA
A;Residues: 1-248 <GAR>
A;Cross-references: UNIPROT:O86282; GB:AE001428; GB:AE001362; NID:g3845316; PIDN:AAC7197
A;Experimental source: clone 3d7
C;Genetics:
A;Gene: PFB0950w

Query Match 44.2%; Score 49.5; DB 2; Length 248;
Best Local Similarity 46.7%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

Qy 2 VCRCLCRGGVCRCLC 16
Db 166 ICRCTC---ICSLC 177

RESULT 13
JC5042
G protein-coupled receptor - barnacle
C;Species: Balanus amphitrite (barnacle)
C;Date: 21-Jan-1997 #sequence_revision 21-Jan-1997 #text_change 09-Jul-2004
C;Accession: JC5042
R;Isoai, A.; Kawahara, H.; Okazaki, Y.; Shizuri, Y.
Gene 175, 95-100, 1996
A;Title: Molecular cloning of a new member of the putative G protein-coupled receptor ge
A;Reference number: JC5042; MUID:97074655; PMID:8917082
A;Accession: JC5042

A;Molecule type: DNA
A;Residues: 1-476 <ISO>
A;Cross-references: UNIPROT:Q9126; DBJ:D78363; NID:g1514430; PIDN:BA11375.1; PID:g1514
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prote
F;34-57/Domain: transmembrane #status predicted <TM1>
F;71-93/Domain: transmembrane #status predicted <TM2>
F;107-129/Domain: transmembrane #status predicted <TM3>
F;128-134/Region: G protein-binding #status predicted
F;148-172/Domain: transmembrane #status predicted <TM4>
F;199-222/Domain: transmembrane #status predicted <TM5>
F;377-398/Domain: transmembrane #status predicted <TM6>
F;407-430/Domain: transmembrane #status predicted <TM7>
F;13.17/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;230,240,250,260/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status
F;355/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predict

Query Match 43.8%; Score 49; DB 2; Length 476;
Best Local Similarity 47.1%; Pred. No. 49;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 VCRCLCRGGVCRCLCR 18
Db 442 LCKVCRRGAMERRFR 458

RESULT 14
G72858
AcOrf-70 protein - Autographa californica nuclear polyhedrosis virus
C;Species: Autographa californica nuclear polyhedrosis virus, ACMPV
A;Note: dsDNA virus
C;Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 09-Jul-2004
C;Accession: G72858
R;Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A;Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.
A;Reference number: A72850; MUID:94303173; PMID:8030224
A;Accession: G72858
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <AVR>
A;Cross-references: UNIPROT:P41470; GB:L22858; NID:g510708; PIDN:AAA66700.1; PID:g559139
C;Genetics:
A;Gene: AcOrf-70

Query Match 42.9%; Score 48; DB 2; Length 290;
Best Local Similarity 52.9%; Pred. No. 47;
Matches 9; Conservative 3; Mismatches 3; Indels 2; Gaps 2;

Qy 4 RCLC-RRGV-CRCLCR 18
Db 228 QCFPRQGYKCEICRR 244

RESULT 15
B70554
Hypothetical protein Rv1145 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: B70554
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: B70554
A;Status: preliminary;
A;Molecule type: DNA
A;Residues: 1-303 <COL>
A;Cross-references: UNIPROT:O06545; GB:Z95584; GB:AL123456; NID:g3261774; PIDN:CAB09033.1
A;Experimental source: strain H37RV

C;Genetics:
A;Gene: RV1145

Query Match 42.9%; Score 48; DB 2; Length 303;
Best Local Similarity 60.0%; Pred. No. 48;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 RCLCRGVRCICLR 18
||| ||| |||
DB 281 RCCFRSPWRCRR 295

EGMSG
epidermal growth factor precursor - mouse
N;Alternate names: urogastrone precursor
C;Species: Mus musculus (house mouse)
C;Date: 30-Nov-1980 #sequence_revision 11-Aug-1983 #text_change 09-Jul-2004
C;Accession: A94272; A93304; A92118; A01387
R;Scott, J.; Urdea, M.; Quiroga, M.; Sanchez-Bescador, R.; Fong, N.; Selby, M.; Rutter,
Science 211, 236-240, 1983
A;Title: Structure of a mouse submaxillary messenger RNA encoding epidermal growth facto
A;Reference number: A94272; MUID:83223630; PMID:6602382
A;Accession: A94272
A;Molecule type: mRNA
A;Residues: 1-1217 <SCO>
A;Cross-references: UNIPROT:P01132; GB:J00380; NID:g192993; PIDN:AAA37539.1; PID:G309210
R;Gray, A.; Dull, T.; Ullrich, A.
Nature 303, 722-725, 1983
A;Title: Nucleotide sequence of epidermal growth factor cDNA predicts a 128,000-molecula
A;Reference number: A93304; MUID:83219309; PMID:6304537
A;Accession: A93304
A;Molecule type: mRNA
A;Residues: 1-789, 'Y', 791-1047, 'S', 1049-1168 <GRA>
A;Cross-references: GB:J00380
A;Note: The sequence shown by these authors differs from residues 1134-1168 due to an in
uence of Scott et al.)
R;Savage Jr., C.R.; Inagami, T.; Cohen, S.
J. Biol. Chem. 247, 7612-7621, 1972
A;Title: The primary structure of epidermal growth factor.
A;Reference number: A92118; MUID:73048516; PMID:4636327
A;Accession: A92118
A;Molecule type: protein
A;Residues: 977-1029 <SAV>
A;Note: Residues 1024-1029 are not required for full biological activity in vivo
R;Savage Jr., C.R.; Hash, J.H.; Cohen, S.
J. Biol. Chem. 248, 7669-7672, 1973
A;Title: Epidermal growth factor. Location of disulfide bonds.
A;Reference number: A92144; MUID:74025498; PMID:4750422
A;Contents: annotation; disulfide bonds
C;Comment: Epidermal growth factor (EGF) stimulates the proliferation and differentiati
gastrointestinal cell proliferation.
C;Comment: EGF is released in the pancreas, small intestine, mammary gland, and (in some
C;Comment: The EGF precursor is found in kidney as a receptor-like membrane-bound protein
C;Comment: The active growth factor from this submaxillary gland protein stimulates the

C;Superfamily: epidermal growth factor precursor; EGF homology; LDL receptor YWTD-contai
C;Keywords: duplication; growth factor; tandem repeat; transmembrane protein
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-1029/Product: epidermal growth factor proprotein, membrane-bound form #status predi
F;29-1028/Domain: extracellular #status predicted <EXT>
F;50-485/Region: EGF precursor long repeat
F;53-92/Domain: LDL receptor YWTD-containing repeat homology <YW01>
F;93-134/Domain: LDL receptor YWTD-containing repeat homology <YW02>
F;135-176/Domain: LDL receptor YWTD-containing repeat homology <YW03>
F;177-217/Domain: LDL receptor YWTD-containing repeat homology <YW04>
F;218-262/Domain: LDL receptor YWTD-containing repeat homology <YW05>
F;263-307/Domain: LDL receptor YWTD-containing repeat homology <YW06>
F;324-360/Domain: EGF homology #status atypical <EG1>
F;366-401/Domain: EGF homology <EG2>
F;407-442/Domain: EGF homology <EG3>
F;445-482/Domain: EGF homology <EG4>
F;486-961/Region: EGF precursor long repeat

F;489-529/Domain: LDL receptor YWTD-containing repeat homology <YW07>
F;530-572/Domain: LDL receptor YWTD-containing repeat homology <YW08>
F;573-615/Domain: LDL receptor YWTD-containing repeat homology <YW09>
F;616-659/Domain: LDL receptor YWTD-containing repeat homology <YW10>
F;660-700/Domain: LDL receptor YWTD-containing repeat homology <YW11>
F;701-743/Domain: LDL receptor YWTD-containing repeat homology <YW12>
F;751-786/Domain: EGF homology <EG5>
F;842-875/Domain: EGF homology <EG6>
F;881-917/Domain: EGF homology <EG7>
F;923-958/Domain: EGF homology <EG8>
F;977-1029/Product: epidermal growth factor #status experimental <EGP>
F;982-1018/Domain: EGF homology <EG9>
F;1039-1063/Domain: transmembrane #status predicted <TMM>
F;1064-1217/Domain: intracellular #status predicted <INT>
F;347-360.366-377.373-386.388-401.407-418.414-427.429-442.445-457.453-467.469-482.751-762
tus predicted
F;982-996.990-1007.1009-1018/disulfide bonds: #status experimental

Query Match 42.9%; Score 48; DB 1; Length 1217;
Best Local Similarity 47.4%; Pred. No. 1.2e+02;
Matches 9; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 GVCECLCR--GVCECLCR 17
||| :||| |||
DB 756 GGCEHCICCESLGTARCLCR 774

RESULT 17
A41116
transcription factor AMT1, metal-inducible - yeast (Candida glabrata)
C;Species: Candida glabrata
C;Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 09-Jul-2004
C;Accession: A41116
R;Zhou, P.; Thiele, D.J.
Proc. Natl. Acad. Sci. U.S.A. 88, 6112-6116, 1991
A;Title: Isolation of a metal-activated transcription factor gene from Candida glabrata t
A;Reference number: A41116; MUID:91296771; PMID:2068090
A;Accession: A4116
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-265 <ZHO>
A;Cross-references: UNIPROT:P41772; GB:M69146; NID:g173319; PIDN:AAA35271.1; PID:g173320

Query Match 42.4%; Score 47.5; DB 2; Length 265;
Best Local Similarity 66.7%; Pred. No. 50;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 5 CLCRRG-VCRCLCR 18
||| ||| |||
DB 90 CLCVRGEPCHARR 104

RESULT 18
S08167
Balbiani ring 3 protein - midge (Chironomus tentans)
C;Species: Chironomus tentans
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Aug-2004
C;Accession: S08167
R;Paulsson, G.; Lendahl, U.; Galli, J.; Ericson, C.; Wieslander, L.
J. Mol. Biol. 211, 331-349, 1990
A;Title: The Balbiani ring 3 gene in Chironomus tentans has a diverged repetitive struct
A;Reference number: S08167; MUID:90172404; PMID:1689777
A;Accession: S08167
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-1700 <PAU>
A;Cross-references: UNIPROT:Q03376; GB:X52263; NID:g7057; PIDN:CAA36506.1; PID:g7058
C;Genetics:
A;Gene: BR3
A;Map position: 4

Query Match 42.0%; Score 47; DB 2; Length 1700;
Best Local Similarity 22.6%; Pred. No. 1.9e+02;

Matches 7; Conservative 6; Mismatches 2; Indels 16; Gaps 1;
 Qy 3 CRCLRRG-----VCRCLCR 17
 Db 1245 CRVCPRNMEKPADNCKTKWNDEMCCVCCK 1275
 RESULT 19
 GNNY21
 genome polyprotein - coxsackievirus A21 (strain Coe)
 N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
 polyomerase (EC 2.7.7.48)
 C;Species: coxsackievirus A21
 C;Accession: A33373
 C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004
 R;Hughes, P.J.; North, C.; Minor, P.D.; Stanway, G.
 J. Gen. Virol. 70, 2943-2952, 1989
 A;Title: The complete nucleotide sequence of coxsackievirus A21.
 A;Reference number: A33373; MUID:90063544; PMID:2584950
 A;Accession: A33373
 A;Molecule type: Genomic RNA
 A;Residues: 1-2206 <HUG>
 A;Cross-references: UNIPROT:P22055; GB:D00538; NID:G221147; PIDN:BA00426.1; PID:G221148
 C;Superfamily: poliovirus genome polyprotein
 C;Keywords: coat protein; core protein; genome-linked protein; nucleotidyltransferase; p
 F;1-69/Product: coat protein 1A #status predicted <VP4>
 F;70-341/Product: coat protein 1B #status predicted <VP2>
 F;342-578/Product: coat protein 1C #status predicted <VP3>
 F;579-881/Product: coat protein 1D #status predicted <VP1>
 F;882-1028/Product: core protein 2A #status predicted <PA2>
 F;1029-1125/Product: core protein 2B #status predicted <PB2>
 F;1126-1453/Product: core protein 2C #status predicted <PC2>
 F;1454-1540/Product: protein 3A #status predicted <PA3>
 F;1541-1562/Product: genome-linked protein VPg #status predicted <PB3>
 F;1563-1745/Product: proteinase #status predicted <PC3>
 F;1746-2206/Product: RNA-directed RNA polymerase #status predicted <PD3>
 F;1543/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted
 Query Match 42.0%; Score 47; DB 1; Length 2206;
 Best Local Similarity 52.9%; Pred. No. 2.2e+02;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 Qy 2 VCRCLRRGVCRCRCR 18
 Db 931 IARCSCRCGVYCESRR 947
 RESULT 20
 T42214
 otogelin - mouse
 N;Alternate names: mucin-like extracellular matrix protein
 C;Species: Mus musculus (house mouse)
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Aug-2004
 C;Accession: T42214
 R;Cohen-Salmon, M.; El-Amraoui, A.; Leibovici, M.; Petit, C.
 Proc. Natl. Acad. Sci. U.S.A. 94, 14450-14455, 1997
 A;Title: Otogelin: A glycoprotein specific to the acellular membranes of the inner ear.
 A;Reference number: 222079; MUID:98070772; PMID:9405633
 A;Accession: T42214
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-2910 <COH>
 A;Cross-references: UNIPROT:O55225; EMBL:U96411; NID:G2760883; PID:G2760884; PIDN:AAB965
 A;Experimental source: strain BAUB/c
 A;Note: component of all the acellular membranes of the inner ear
 C;Superfamily: von Willebrand factor type A repeat homology; von Willebrand factor type
 Query Match 42.0%; Score 47; DB 2; Length 2910;
 Best Local Similarity 38.5%; Pred. No. 2.6e+02;
 Matches 10; Conservative 0; Mismatches 2; Indels 14; Gaps 1;
 Qy 7 CRGVCR-----LCRR 18
 Db 7 CRGVCR-----LCRR 18

Db 724 CRDRCRGQPCLCATLAHYARLCRR 749
 RESULT 21
 I48725
 Q300 protein - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C;Accession: I48725
 R;Wagner, S.; Cullmann, G.; Knippers, R.
 J. Virol. 65, 3259-3267, 1991
 A;Title: The Q300 gene: a novel transcription unit induced in simian virus 40-infected ar
 A;Reference number: I48725; MUID:91237845; PMID:1851876
 A;Accession: I48725
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-77 <RES>
 A;Cross-references: UNIPROT:Q02722; EMBL:X52164; NID:G53860; PIDN:CAA36417.1; PID:G53861
 Query Match 41.5%; Score 46.5; DB 2; Length 77;
 Best Local Similarity 53.3%; Pred. No. 30;
 Matches 8; Conservative 2; Mismatches 3; Gaps 1;
 Qy 2 VCRCLRRGVCRCRC 16
 Db 27 VCVCVC---VCVCVC 38
 RESULT 22
 T46871
 C-8 sterol isomerase (EC 5.3.3.-) [imported] - Neurospora crassa
 N;Alternate names: delta 8->delta 7 sterol isomerase
 C;Species: Neurospora crassa
 C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
 C;Accession: I46871
 R;Gilbert, J.; Orbach, M.J.
 submitted to the EMBL Data Library, May 1996
 A;Reference number: 224117
 A;Accession: T46871
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-256 <GIL>
 A;Cross-references: UNIPROT:Q92254; EMBL:U59671; PIDN:AAB09470.1
 A;Experimental source: strain OR23-74-1A
 C;Genetics:
 A;Gene: erg-1
 A;Map position: V
 A;Introns: 145/1
 C;Keywords: intramolecular oxidoreductase; isomerase; sterol biosynthesis
 Query Match 41.5%; Score 46.5; DB 2; Length 256;
 Best Local Similarity 56.7%; Pred. No. 64;
 Matches 8; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
 Qy 3 CRCLRRGVCR 14
 Db 30 CRCSCR---CRC 38
 RESULT 23
 JP0076
 nel protein - chicken
 C;Species: Gallus gallus (chicken)
 C;Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 04-Apr-2004
 C;Accession: A38963; JP0076
 R;Matsushashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.; Taniguchi, S.; Hori, K.
 Dev. Dyn. 203, 212-222, 1995
 A;Title: New gene, nel, encoding a Mr 93K protein with EGF-like repeats is strongly expre
 A;Reference number: A38963; MUID:95383734; PMID:7655083
 A;Accession: A38963
 A;Molecule type: mRNA
 A;Residues: 1-835 <MAT>
 A;Cross-references: DDBJ:D45365

A;Experimental source: 9-day embryo
 R;Matsushashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.; Taniguchi, S.; Hori, K.
 submitted to JIPID, January 1995
 A;Description: A new gene, nel, encoding a Mr 93K protein with EGF-like repeats is stron
 A;Reference number: JP0076
 A;Accession: JP0076

A;Molecule type: mRNA
 A;Residues: 1-835 <MA2>
 A;Cross-references: DDBJ:D45365
 A;Experimental source: 9-day embryo
 F;273-333/Domain: von Willebrand factor type C repeat homology <VWC>
 F;395-592/Region: EGF-like repeats
 F;444-480/Domain: EGF homology <EGF1>
 F;486-521/Domain: EGF homology <EGF>
 F;525-552/Domain: EGF homology <EGF2>

Query Match 41.5%; Score 46.5; DB 2; Length 835;
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;
 Matches 9; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

Qy 2 VCRG-----LCRRGVCR 14
 ||| ||| ||| ||| |||
 Db 662 VCSCQGVYVCRMVVDC 679

RESULT 24
 S29214
 neurotoxin Tx2 - spider (Phoneutria nigriventer)
 C;Species: Phoneutria nigriventer
 C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
 C;Accession: S29214
 R;do Nascimento Cordeiro, M.; Ribeiro Diniz, C.; do Carmo Valentim, A.; von Eickstedt, V.
 FEBS Lett. 310, 153-156, 1992
 A;Title: The purification and amino acid sequences of four Tx2 neurotoxins from the vena
 A;Reference number: S29214; PMID:1397265
 A;Accession: S29214
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-53 <COR>
 C;Superfamily: curtaxotoxin

Query Match 41.1%; Score 46; DB 2; Length 53;
 Best Local Similarity 34.6%; Pred. No. 27;
 Matches 9; Conservative 2; Mismatches 5; Indels 10; Gaps 1;

Qy 3 CRLCRGVGVC-----RCLCR 18
 ||| ||| ||| ||| |||
 Db 14 CDCGGERGCVGCVSVEGKYRCICRQ 39

RESULT 25
 LNP61
 pulmonary surfactant protein 9K form - pig
 N;Alternate names: low molecular mass surfactant protein type 1
 C;Species: Sus scrofa domestica (domestic pig)
 C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
 C;Accession: S00363
 R;Curstedt, T.; Johansson, J.; Barros-Soederling, J.; Robertson, B.; Nilsson, G.; Westbe
 Eur. J. Biochem. 172, 521-525, 1988
 A;Title: Low-molecular-mass surfactant protein type 1. The primary structure of a hydroph
 A;Reference number: S00363; PMID:88166729; PMID:3350011
 A;Accession: S00363
 A;Molecule type: protein
 A;Residues: 1-79 <CUR>

A;Cross-references: UNIPROT:P15782
 C;Comment: Pulmonary surfactant protein is a phospholipid-protein complex, which reduces
 C;Superfamily: pulmonary surfactant protein B; saposin repeat homology
 C;Keywords: alveolar proteinosis; gaseous exchange; lipoprotein; lung; pulmonary surfact
 F;1-79/Domain: saposin repeat homology <SAP>

Query Match 41.1%; Score 46; DB 1; Length 79;
 Best Local Similarity 46.7%; Pred. No. 35;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GVCRLCRRGVCRCL 15
 ||| ||| ||| ||| |||
 Db 44 GICQCLAERYIVICL 58

RESULT 26
 C37057
 fibronectin receptor beta chain - guinea pig (fragment)
 C;Species: Cavia porcellus (guinea pig)
 C;Date: 15-Feb-1991 #sequence_revision 15-Feb-1991 #text_change 16-Aug-2004
 C;Accession: C37057
 R;Sheppard, D.; Rozzo, C.; Starr, L.; Quaranta, V.; Erle, D.J.; Pytela, R.
 J. Biol. Chem. 265, 11502-11507, 1990
 A;Title: Complete amino acid sequence of a novel integrin beta subunit (beta6) identified
 A;Reference number: A37057; PMID:90307659; PMID:2365883
 A;Accession: C37057
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-94 <SHE>
 A;Cross-references: GB:J05522
 C;Superfamily: Integrin beta chain; laminin-type EGF-like homology
 C;Keywords: cell adhesion; duplication; heterodimer; membrane protein

Query Match 41.1%; Score 46; DB 2; Length 94;
 Best Local Similarity 57.1%; Pred. No. 39;
 Matches 8; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

Qy 1 GVCRLCRRGVCR 14
 ||| ||| ||| ||| |||
 Db 76 GVCEC-----GVCKC 85

RESULT 27
 D72481
 hypothetical protein APE2492 - Aeropyrum pernix (strain K1)
 C;Species: Aeropyrum pernix
 C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C;Accession: D72481
 R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaha
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki
 DNA Res. 6, 83-101, 1999
 A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A;Reference number: A72450; PMID:99310339; PMID:10382966
 A;Accession: D72481
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-128 <KAW>
 A;Cross-references: UNIPROT:Q9Y8Z2; DDBJ:AP000064; NID:G5105945; PIDN:BAA81508.1; PID:di
 A;Experimental source: strain K1
 C;Genetics:
 A;Gene: APE2492
 C;Superfamily: Aeropyrum pernix hypothetical protein APE2492

Query Match 41.1%; Score 46; DB 2; Length 128;
 Best Local Similarity 47.6%; Pred. No. 47;
 Matches 10; Conservative 2; Mismatches 5; Indels 4; Gaps 1;

Qy 1 GVCVCR-----CLCRRGVCRCLCR 17
 ||| ||| ||| ||| |||
 Db 41 GNCRESILNLQQLGLCRFCR 61

RESULT 28
 A25854
 chloramphenicol resistance protein - Escherichia coli plasmid R26
 C;Species: Escherichia coli
 C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jul-2004
 C;Accession: A25854
 R;Dorman, C.J.; Foster, T.J.; Shaw, W.V.
 Gene 41, 349-353, 1986
 A;Title: Nucleotide sequence of the R26 chloramphenicol resistance determinant and identi
 A;Reference number: A25854; PMID:86221720; PMID:3011609


```

A:Accession: A25854
A:Molecule type: DNA
A:Residues: 1-302 <DOR>
A:Cross-references: UNIPROT:P12056; GB:M22614; NID:g151802; PIDN:AAA26079.1; PID:g151803
C:Genetics:
A:Gene: cml
A:Genome: plasmid
C:Superfamily: bicyclomycin resistance protein
C:Keywords: antibiotic resistance

Query Match 41.1%; Score 46; DB 2; Length 302;
Best Local Similarity 57.1%; Pred. No. 82;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 CLCRRGVCRCLCR 18
DB 35 CVSRFGHRCICR 48

RESULT 29
H90942
probable diogenase beta subunit [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H90942
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90942
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <HAY>
A:Cross-references: UNIPROT:Q8XDR8; GB:BA000007; PIDN:HAB35935.1; PID:g13361979; GSPDB:G
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC62512
C:Superfamily: phthalate dioxygenase reductase; cytochrome-b5 reductase homology; ferred

Query Match 41.1%; Score 46; DB 2; Length 321;
Best Local Similarity 87.5%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CLCRRGVC 12
DB 268 CLCREGVC 275

RESULT 30
C64941
probable diogenase beta chain - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 16-Aug-2004
C:Accession: C64941
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: C64941
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-321 <BLAT>
A:Cross-references: UNIPROT:P76254; GB:AE000274; GB:U00096; NID:g1788089; PIDN:AAC74873
A:Experimental source: strain K-12, substrain MG1655
C:Superfamily: Phthalate dioxygenase reductase; cytochrome-b5 reductase homology; ferred
F;11-225/Domain: cytochrome-b5 reductase homology <CBR>
F;254-309/Domain: ferredoxin [2Fe-2S] homology <FER>

Query Match 41.1%; Score 46; DB 2; Length 321;
Best Local Similarity 87.5%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CLCRRGVC 12
DB 268 CLCREGVC 275

RESULT 31
D85791
probable diogenase beta subunit yeast [imported] - Escherichia coli (strain O157:H7, subst
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Aug-2004
C:Accession: D85791
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85791
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <STO>
A:Cross-references: UNIPROT:Q8XDR8; GB:AE005174; NID:g12515843; PIDN:AAG56792.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yeast
C:Superfamily: phthalate dioxygenase reductase; cytochrome-b5 reductase homology; ferred

Query Match 41.1%; Score 46; DB 2; Length 321;
Best Local Similarity 87.5%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CLCRRGVC 12
DB 268 CLCREGVC 275

RESULT 32
C82426
cobryic acid synthase VCA0727 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: C82426
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, B.; Ermolaeva, M.D.; Vamathevan, J.; Basu, S.; Qin, H.; Dragoi, I.; Sellers, P.
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: C82426
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-484 <HEI>
A:Cross-references: UNIPROT:Q9KLL6; GB:AE004401; GB:AE003853; NID:g9658132; PIDN:AAF9662
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0727
A:Map position: 2
C:Superfamily: probable cobryic acid synthase

Query Match 41.1%; Score 46; DB 2; Length 484;
Best Local Similarity 57.1%; Pred. No. 11e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GVCRCCLCRRGVCRC 14
DB 22 GLCRVLARRGIQVC 35

RESULT 33
EGRT
epidermal growth factor precursor - rat
N:Alternate names: urogastrone precursor
C:Species: Rattus norvegicus (Norway rat)

```


C>Date: 31-Dec-1988 #sequence revision 14-Aug-1998 #text change 09-Jul-2004
C/Accession: I52995; S05074; S01974; A25425; S18419; S08288
R/Saggi, S.J.; Safirstein, R.; Price, P.M.
DNA Cell Biol. 11, 481-487, 1992
A/Title: Cloning and Sequencing of the Rat Preproepidermal Growth Factor cDNA: Comparison
A/Reference number: I52995; MUID:92398779; PMID:1524680
A/Accession: I52995
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1133 <RES>
A/Cross-references: UNIPROT:P07522; EMBL:U04842; NID:G440236; PIDN:AAB60436.1; PID:G4402
R/Simpson, R.
submitted to the EMBL Data Library, August 1988
A/Reference number: S05074
A/Accession: S05074
A/Molecule type: mRNA
A/Residues: 'W', 966, 'RWL', 970-1023, 'NW', 1026-1108, 'SGAGVSSGPWFVVLE', 1126, 'HQ' <SIM>
A/Cross-references: EMBL:X12748
R/Dorow, D.S.; Simpson, R.J.
Nucleic Acids Res. 16, 9338, 1988
A/Title: Cloning and sequence analysis of a cDNA for rat epidermal growth factor.
A/Reference number: S01974; MUID:89016634; PMID:3262867
A/Accession: S01974
A/Molecule type: mRNA
A/Residues: 'W', 966, 'RWL', 970-1023, 'NW', 1026-1108 <DOR>
A/Cross-references: EMBL:X12748
R/Simpson, R.J.; Smith, J.A.; Moritz, R.L.; O'Hare, M.J.; Rudland, P.S.; Morrison, J.R.;
Eur. J. Biochem. 153, 629-637, 1985
A/Title: Rat epidermal growth factor: complete amino acid sequence.
A/Reference number: A25425; MUID:86081810; PMID:3000782
A/Accession: A25425
A/Molecule type: protein
R/Nishi, N.; Shimizu, C.; Okutani, T.; Kagawa, Y.; Takasuga, H.; Suno, M.; Wada, F.
Biochim. Biophys. Acta 1095, 268-275, 1991
A/Title: Rat prostatic growth factors: purification and characterization of high and low
A/Reference number: S18419; MUID:92069070; PMID:1958699
A/Accession: S18419
A/Status: preliminary
A/Molecule type: protein
A/Residues: 974-1021 <NIS>
R/Nexo, E.; Jorgensen, P.E.; Thim, L.; Roepstorff, P.
Biochim. Biophys. Acta 1037, 388-393, 1990
A/Title: Purification and characterization of a low and a high molecular weight form of
A/Reference number: S08288; MUID:90181442; PMID:2310752
A/Accession: S08288
A/Molecule type: protein
A/Residues: 974-1024 <NEX>
C/Comment: Epidermal growth factor (EGF) stimulates the proliferation and differentiation
gastrointestinal cell proliferation.
C/Comment: EGF is released in the pancreas, small intestine, mammary gland, and (in some
C/Comment: The EGF precursor is found in kidney as a receptor-like membrane-bound protein
C/Superfamily: epidermal growth factor precursor; EGF homology; LDL receptor YWTD-contai
C/Keywords: duplication; growth factor; mitogen; tandem repeat; transmembrane protein
P:1-21/Domain: signal sequence #status predicted <SIG>
P:22-1133/Product: epidermal growth factor proprotein, membrane-bound form #status predi
F:22-1035/Domain: extracellular #status predicted <EXT>
F:44-480/Region: EGF precursor long repeat <LR1>
F:47-86/Domain: LDL receptor YWTD-containing repeat homology <YW01>
F:87-128/Domain: LDL receptor YWTD-containing repeat homology <YW02>
F:129-170/Domain: LDL receptor YWTD-containing repeat homology <YW03>
F:171-212/Domain: LDL receptor YWTD-containing repeat homology <YW04>
F:213-257/Domain: LDL receptor YWTD-containing repeat homology <YW05>
F:258-302/Domain: LDL receptor YWTD-containing repeat homology <YW06>
F:319-355/Domain: LDL receptor YWTD-containing repeat homology <YW07>
F:361-396/Domain: EGF homology <EG2>
F:402-437/Domain: EGF homology <EG3>
F:440-477/Domain: EGF homology <EG4>
F:482-558/Region: EGF precursor long repeat <LR2>
F:485-525/Domain: LDL receptor YWTD-containing repeat homology <YW07>
F:526-568/Domain: LDL receptor YWTD-containing repeat homology <YW08>
F:569-611/Domain: LDL receptor YWTD-containing repeat homology <YW09>
F:612-655/Domain: LDL receptor YWTD-containing repeat homology <YW10>

F:656-696/Domain: LDL receptor YWTD-containing repeat homology <YW11>
F:697-739/Domain: LDL receptor YWTD-containing repeat homology <YW12>
F:747-782/Domain: EGF homology <EG5>
F:839-872/Domain: EGF homology <EG6>
F:878-914/Domain: EGF homology <EG7>
F:920-955/Domain: EGF homology <EG8>
F:974-1024/Product: epidermal growth factor #status experimental <MAT>
F:979-1015/Domain: EGF homology <EG9>
F:1036-1060/Domain: transmembrane #status predicted <TMM>
F:1061-1133/Domain: intracellular #status predicted <INT>
F:1342-3553/1-372,368-381,383-396,402-413,409-422,424-437,440-452,448-462,464-477,747-758
-1015/disulfide bonds: #status predicted
Query Match 41.1%; Score 46; DB 1; Length 1133;
Best Local Similarity 37.8%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 6; Gaps 1;
QY 1 GVCRCLCRGVC-----CLCRR 18
Db 763 GTAQCLCRGFKVAPDGMCLTRK 786
RESULT 34
JH0675
restrictin precursor - chicken
C/Species: Gallus gallus (chicken)
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
C/Accession: JH0675; PS0385; S23254
R/Noerenberg, U.; Wille, H.; Wolff, J.M.; Frank, R.; Rathjen, F.G.
Neuron 8, 849-863, 1992
A/Title: The chicken neural extracellular matrix molecule restrictin: similarity with EGF
A/Reference number: JH0675; MUID:92265298; PMID:1375037
A/Accession: JH0675
A/Molecule type: mRNA
A/Residues: 1-1353 <NOE>
A/Cross-references: UNIPROT:Q00546; GB:X64649; NID:G63613; PIDN:CAA45920.1; PID:G63614
A/Experimental source: brain
A/Accession: PS0385
A/Molecule type: protein
A/Residues: 579-586;827-840 <NOE1>
C/Comment: This protein is a neural extracellular matrix protein implicated in neural cel
C/Superfamily: restrictin; EGF homology; fibronogen beta/gamma homology; fibronectin type
C/Keywords: calcium binding; cell adhesion; duplication; extracellular matrix; glycoprote
F:1-33/Domain: signal sequence #status predicted <SIG>
F:34-133/Product: restrictin #status predicted <MAT>
F:203-229/Domain: EGF homology <EG1>
F:234-260/Domain: EGF homology <EG2>
F:265-291/Domain: EGF homology <EG3>
F:296-322/Domain: EGF homology <EG4>
F:324-405/Domain: fibronectin type III repeat homology <FN1>
F:413-494/Domain: fibronectin type III repeat homology <FN2>
F:502-584/Domain: fibronectin type III repeat homology <FN3>
F:592-676/Domain: fibronectin type III repeat homology <FN4>
F:684-764/Domain: fibronectin type III repeat homology <FN5>
F:772-853/Domain: fibronectin type III repeat homology <FN6>
F:861-941/Domain: fibronectin type III repeat homology <FN7>
F:949-1027/Domain: fibronectin type III repeat homology <FN8>
F:1035-1115/Domain: fibronectin type III repeat homology <FN9>
F:1130-1338/Domain: fibronogen beta/gamma homology <FBG>
F:1272-1286/Region: calcium binding #status predicted
F:53,197,277,391,469,580,734,790,960,1031,1041,1256,1342/Binding site: carbohydrate (Asn)

Query Match 41.1%; Score 46; DB 1; Length 1353;
Best Local Similarity 27.8%; Pred. No. 2.1e+02;
Matches 10; Conservative 3; Mismatches 1; Indels 22; Gaps 2;
QY 3 CRCLC-----RRGVC---RCIC 16
Db 185 CRICSEGAGNSCNEPRCPRGCSRGVLCGQVC 220
RESULT 35
S07405

protease inhibitor (Bowman-Birk) C-II - soybean
 C/Species: Glycine max (soybean)
 C/Date: 02-Dec-1993 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
 C/Accession: S07405; S29608; S40113
 R/Joudrier, P.E.; Foard, D.E.; Floener, L.A.; Larkins, B.A.
 Plant Mol. Biol. 10, 35-42, 1987
 A/Title: Isolation and sequence of cDNA encoding the soybean protease inhibitors PI IV a
 A/Reference number: S07405
 A/Accession: S07405
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-83 <BA>
 A/Cross-references: UNIPROT:P01063; EMBL:M20732; NID:G169944; PIDN:AAA33953.1; PID:G16994
 A/Note: the sequences of codons 8-13 and 14-27 are interchanged in the authors' translat
 R/Baek, J.M.; Kim, S.I.
 submitted to the EMBL Data Library, October 1992
 A/Description: Nucleotide sequence of a cDNA encoding the soybean Bowman-Birk proteinase
 A/Reference number: S29559
 A/Accession: S29559
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-83 <BA>
 A/Cross-references: EMBL:X69705; NID:G18567; PIDN:CAA48656.1; PID:G18568
 R/Giordano, A.; Dalledonne, M.; Fogher, C.; Marchetti, S.
 submitted to the EMBL Data Library, December 1993
 A/Description: Nucleotide sequence encoding a soybean C-II proteinase inhibitor.
 A/Reference number: S40113
 A/Accession: S40113
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-83 <GI>
 A/Cross-references: EMBL:X76727; NID:G436413; PIDN:CAA54144.1; PID:G436414
 C/Superfamily: Bowman-Birk proteinase inhibitor; Bowman-Birk inhibitor repeat homology
 C/Keywords: serine proteinase inhibitor
 F:22-48/Domain: Bowman-Birk inhibitor repeat homology <BB1>
 F:43-74/Domain: Bowman-Birk inhibitor repeat homology <BB12>
 Query Match 40.6%; Score 45.5; DB 2; Length 83;
 Best Local Similarity 60.0%; Pred. No. 41;
 Matches 9; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
 QY 4 RCLCRR---GVCRCCL 15
 DB 51 RCACTRSMPCQCRCCL 65
 RESULT 36
 JC2225
 Bowman-Birk proteinase isoinhibitor C-II precursor (clone pB24) - soybean
 C/Species: Glycine max (soybean)
 C/Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 18-Aug-2000
 C/Accession: JC2225
 R/Baek, J.M.; Song, J.C.; Choi, Y.D.; Kim, S.I.
 Biosci. Biotechnol. Biochem. 58, 843-846, 1994
 A/Title: Nucleotide sequence homology of cDNAs encoding soybean Bowman-Birk type protein
 A/Reference number: JC2224; MUID:94289861; PMID:7764974
 A/Accession: JC2225
 A/Molecule type: mRNA
 A/Residues: 1-94 <BAE>
 C/Comment: This protein regulates endogeneous proteinase during germination, stores sulf
 C/Genetics:
 A/Start codon: GTT
 C/Superfamily: Bowman-Birk proteinase inhibitor; Bowman-Birk inhibitor repeat homology
 C/Keywords: duplication; seed; serine proteinase inhibitor
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-94/Product: Bowman-Birk proteinase isoinhibitor C-II #status predicted <MAT>
 F:33-59/Domain: Bowman-Birk inhibitor repeat homology <BB1>
 F:60-85/Domain: Bowman-Birk inhibitor repeat homology <BB12>
 Query Match 40.6%; Score 45.5; DB 2; Length 94;
 Best Local Similarity 60.0%; Pred. No. 45;
 Matches 9; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 4 RCLCRR---GVCRCCL 15
 DB 62 RCACTRSMPCQCRCCL 76
 RESULT 37
 TISIC2
 proteinase inhibitor (Bowman-Birk) C-II precursor - soybean
 C/Species: Glycine max (soybean)
 C/Date: 24-Apr-1984 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
 C/Accession: A22636; A01302
 R/Hammond, R.W.; Foard, D.E.; Larkins, B.A.
 J. Biol. Chem. 259, 9883-9890, 1984
 A/Title: Molecular cloning and analysis of a gene coding for the Bowman-Birk protease int
 A/Reference number: A92489; MUID:84264652; PMID:6086657
 A/Contents: annotation
 A/Note: the sequence has been revised in reference A92540
 R/Hammond, R.W.; Foard, D.E.; Larkins, B.A.
 J. Biol. Chem. 260, 7806, 1985
 A/Reference number: A92540
 A/Contents: erratum
 A/Accession: A22636
 A/Molecule type: DNA
 A/Residues: 1-103 <HAM>
 A/Cross-references: UNIPROT:P01063; GB:K01967; NID:G169942; PIDN:AAA33952.1; PID:G169943
 R/Odani, S.; Ikemasa, T.
 J. Biochem. 82, 1523-1531, 1977
 A/Title: Studies on soybean trypsin inhibitors. XI. Complete amino acid sequence of a soy
 A/Reference number: A01302; MUID:78087480; PMID:599141
 A/Accession: A01302
 A/Molecule type: protein
 A/Residues: 28-103 <ODA>
 C/Comment: This protein inhibits trypsin, elastase, and chymotrypsin. The site of interac
 C/Superfamily: Bowman-Birk proteinase inhibitor; Bowman-Birk inhibitor repeat homology
 C/Keywords: duplication; seed; serine proteinase inhibitor
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-103/Product: proteinase inhibitor (Bowman-Birk) C-II #status experimental <MAT>
 F:42-68/Domain: Bowman-Birk inhibitor repeat homology <BB1>
 F:69-94/Domain: Bowman-Birk inhibitor repeat homology <BB2>
 F:41-95, 42-57, 45-91, 47-55, 65-72, 69-84, 74-82/Disulfide bonds: #status predicted
 F:49/Inhibitory site: Ala (elastase) #status predicted
 F:76/Inhibitory site: Arg (trypsin) #status experimental
 Query Match 40.6%; Score 45.5; DB 1; Length 103;
 Best Local Similarity 60.0%; Pred. No. 47;
 Matches 9; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
 QY 4 RCLCRR---GVCRCCL 15
 DB 71 RCACTRSMPCQCRCCL 85
 RESULT 38
 T28784
 hypochemical protein C41D11.3 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T28784
 R/Gattung, S.; Maggi, L.
 submitted to the EMBL Data Library, May 1997
 A/Description: The sequence of C. elegans cosmid C41D11.
 A/Reference number: Z20522
 A/Accession: T28784
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-532 <GAT>
 A/Cross-references: UNIPROT:O01972; EMBL:AF003740; PIDN:AAC48138.1; GSPDB:GN00019; CESP:
 A/Experimental source: strain Bristol N2; clone C41D11
 C/Genetics:
 A/Gene: CESP:C41D11.3
 A/Map position: 1
 A/Introns: 62/3; 172/1; 403/3; 478/3
 C/Superfamily: Caenorhabditis elegans hypothetical protein C41D11.3

Query Match 40.6%; Score 45.5; DB 2; Length 532;
Best Local Similarity 47.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 3 CRCLCRRGVC---RCLC 16
| | | | | | | |
Db 298 CGCSCENGVCLEPQC 314

RESULT 39
S02186
hypothetical protein SP5 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004
C;Accession: S02186
R;Nishimatsu, S.I.; Murakami, K.; Mitsui, Y.; Ishida, N.
Nucleic Acids Res. 16, 11831-11832, 1988
A;Title: Mouse spleen derived cDNA clones containing per repeat sequence.
A;Reference number: S02186; MUID:89098345; PMID:3267239
A;Accession: S02186
A;Molecule type: mRNA
A;Residues: 1-95 <NIS>
A;Cross-references: UNIPROT:PI5974; EMBL:X12806; NID:G53649; PIDN:CAB42649.1; PID:G48037

Query Match 40.2%; Score 45; DB 2; Length 95;
Best Local Similarity 52.6%; Pred. No. 51;
Matches 10; Conservative 1; Mismatches 4; Indels 4; Gaps 2;

QY 1 GVCRCCLCRGVCRC---LC 16
| | | | | | | |
Db 8 GVCLCVC-FVCMCVHLC 25

RESULT 40
AE0304
probable dioxxygenase beta chain YPO2492 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 16-Aug-2004
C;Accession: AE0304
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AE0304
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-321 <KUR>
A;Cross-references: UNIPROT:Q8ZDR4; GB:AL590842; PIDN:CAC91297.1; PID:G15980486; GSPDB:Q
C;Genetics:
A;Gene: YPO2492
C;Superfamily: Phthalate dioxxygenase reductase; cytochrome-b5 reductase homology; ferred

Query Match 40.2%; Score 45; DB 2; Length 321;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 CLCRRGVC 12
| | | | | | | |
Db 268 CLCREGIC 275

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 15:13:56 ; Search time 128.5 seconds
(without alignments)
80.597 Million cell updates/sec

Title: US-10-009-317A-33

Perfect score: 112
Sequence: 1 GVCRLCRGVCRCLRR 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02:*

1: uniprot_prot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 60.5 | 54.0 | 243 | Q6ZMP3 | Q6zmp3 homo sapien |
| 2 | 60.5 | 54.0 | 243 | BAD18682 | BAD18682 homo sapi |
| 3 | 60 | 53.6 | 76 | P82271 | P82271 macaca mula |
| 4 | 60 | 53.6 | 168 | Q9D9I2 | Q9d9i2 mus musculus |
| 5 | 58 | 51.8 | 168 | Q6P8T4 | Q6p8t4 mus musculus |
| 6 | 58 | 51.8 | 168 | Q8CH20 | Q8ch20 mus musculus |
| 7 | 58 | 51.8 | 168 | AAH61079 | AAh61079 mus muscu |
| 8 | 58 | 51.8 | 173 | Q9D4K2 | Q9d4k2 mus musculus |
| 9 | 58 | 51.8 | 188 | Q18238 | Q18238 caenorhabdi |
| 10 | 57 | 50.9 | 164 | Q22048 | Q22048 caenorhabdi |
| 11 | 57 | 50.9 | 166 | Q95QY1 | Q95qy1 caenorhabdi |
| 12 | 57 | 50.9 | 197 | Q17641 | Q17641 caenorhabdi |
| 13 | 57 | 50.9 | 274 | Q949G1 | Q949g1 oryza sativ |
| 14 | 56.5 | 50.4 | 190 | Q9UI23 | Q9ui23 homo sapien |
| 15 | 56.5 | 50.4 | 413 | Q69566 | Q69566 human herpe |
| 16 | 55.5 | 49.6 | 512 | Q6RY99 | Q6ry99 rattus norv |
| 17 | 55.5 | 49.6 | 512 | AAZ24072 | AAz24072 rattus no |
| 18 | 55 | 49.1 | 201 | Q6ZQS2 | Q6zqs2 homo sapien |
| 19 | 55 | 49.1 | 201 | BAC87611 | BAC87611 homo sapi |
| 20 | 54.5 | 48.7 | 212 | Q7EDW6 | Q7edw6 anopheles g |
| 21 | 54 | 48.2 | 212 | Q7YVW7 | Q7yvw7 caenorhabdi |
| 22 | 54 | 48.2 | 212 | Q8IHD6 | Q8ihd6 drosophila |
| 23 | 54 | 48.2 | 905 | Q8IMJ2 | Q8imj2 drosophila |
| 24 | 53.5 | 47.8 | 379 | Q7SKV0 | Q7skv0 brachydanio |
| 25 | 53.5 | 47.8 | 418 | O57352 | O57352 coturnix co |
| 26 | 53.5 | 47.8 | 419 | 1 VEGC_HUMAN | P49767 homo sapien |
| 27 | 53.5 | 47.8 | 419 | Q6FHS9 | Q6fhs9 homo sapien |
| 28 | 53.5 | 47.8 | 419 | AAH63685 | AAh63685 homo sapi |
| 29 | 53 | 47.3 | 307 | Q9H9U3 | Q9h9u3 homo sapien |
| 30 | 53 | 47.3 | 881 | Q9W0A0 | Q9w0a0 drosophila |
| 31 | 53 | 47.3 | 1823 | Q7PRP5 | Q7prp5 anopheles g |

32 52.5 46.9 101 2 Q7YU02
33 52.5 46.9 118 2 Q8C9N2
34 52.5 46.9 133 2 Q6R5G9
35 52.5 46.9 133 2 AAR87784
36 52.5 46.9 146 2 Q8BPC0
37 52.5 46.9 160 2 Q9H654
38 52.5 46.9 462 2 Q6GQP2
39 52 46.4 152 2 Q9XVX3
40 52 46.4 1506 2 Q8IRL0
41 52 46.4 23015 2 Q8IQ18
42 52 46.4 23015 2 AAN10358
43 51.5 46.0 83 2 Q7PG24
44 51.5 46.0 107 2 Q8BM14
45 51.5 46.0 134 2 Q8N962

ALIGNMENTS

RESULT 1
Q6ZMP3 PRELIMINARY; PRT; 243 AA.
AC Q6ZMP3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ16784.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Torque;

RA Tanigami A., Fujiwara T., Shibahara T., Goto Y., Hirao M., Shimizu F.,
RA Wakebe H., Ono T., Hishigaki H., Watanabe T., Ozaki K., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA Masuho Y., Nagai K., Isogai T.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
EMBL; AK131548; BAD18682.1; -
DR GO; GO:0016301; F-kinase activity; IEA.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR002219; DAG_PE-Bind.
DR Pfam; PF00130; C1_1; 1.
DR PRINTS; PR00008; DAGPEDOMAIN.
DR SMART; SM00109; C1; 1.
DR PROSITE; PSS0081; DAG_PE_BIND_DOM_2; 1.
KW Kinase.
SQ SEQUENCE 243 AA; 27566 MW; 23D41825EB9F782D CRC64;

Query Match 54.0%; Score 60.5; DB 2; Length 243;
Best Local Similarity 66.7%; Pred. No. 3.4;
Matches 10; Conservative 2; Mismatches 0; Indels 3; Gaps 1;

QY 2 VCRCLCRGVCRCLC 16
||:|||||:|||||
Db 177 VCQCLC---VCQCLC 188

RESULT 2
BAD18682 PRELIMINARY; PRT; 243 AA.
ID BAD18682
AC BAD18682;
DT 12-MAY-2004 (TrEMBLrel. 27, Created)
DT 12-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 12-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE CDNA FLJ16784 fis, clone CTONG2003764, highly similar to Protein
DE kinase C, delta type (EC 2.7.1.-).
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Tongue;
 RA Tanigami A., Fujiwara T., Shibahara T., Goto Y., Hirao M., Shimizu F.,
 RA Wakebe H., Ono T., Hishigaki H., Watanabe T., Ozaki K., Sugiyama T.,
 RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Takakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
 RA Masuho Y., Nagai K., Isogai T.;
 RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AK131548; BAD18682.1;
 DR EMBL; AF194156; AAF07923.1;
 KW Kinase.
 SQ SEQUENCE 243 AA; 27566 MW; 23D41825EB9F782D CRC64;
 Query Match 54.0%; Score 60.5; DB 2; Length 243;
 Best Local Similarity 66.7%; Pred. No. 3.4;
 Matches 10; Conservative 2; Mismatches 0; Indels 3; Gaps 1;
 QY 2 VRCCLCRGVCRCLC 16
 DB 177 VCQCCLC---VCQCCLC 188
 RESULT 3
 P82271
 ID P82271 PRELIMINARY; PRT; 76 AA.
 AC P82271;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Theta defensin-1, subunit B precursor (RTD-1).
 GN Name=RTD1A;
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheidae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.; SEQUENCE OF 65-73, AND MASS SPECTROMETRY.
 RC TISSUE=Bone marrow, and Leukocyte;
 RX MEDLINE=99453140; PubMed=10521339;
 RA Tang Y.-Q., Yuan J., Osapay G., Tran D., Miller C.J.,
 RA Ouellette A.J., Selsted M.E.;
 RT "A Cyclic Antimicrobial Peptide Produced in Primate Leukocytes by the
 RT Ligation of Two Truncated alpha-Defensins.";
 RL Science 286:498-502(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Zhao C., Nguyen T., Lehrer R.I.;
 RT "CDNA cloning of three alpha-defensins and three defensins from
 RT rhesus monkey bone marrow.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Microbicidal activity against Gram-positive bacteria
 CC and E.coli ML35 and fungi C.albicans and C.neoformans in vitro.
 CC -1- SUBUNIT: Heterodimer of subunit A and subunit B linked by a
 CC disulfid bond at position 66 forming a cyclic RTD-1.
 CC -1- TISSUE SPECIFICITY: Bone marrow; promyelocytes, myelocytes and
 CC mature neutrophils and monocytes.
 CC -1- DEVELOPMENTAL STAGE: Expression and peptide accumulation starts
 CC early during granulocyte myelopoiesis.
 CC -1- MASS SPECTROMETRY: MW=2082.0; METHOD=NALDI.
 CC -1- MISCELLANEOUS: The determined pI of this protein is greater than
 CC 12.
 CC -1- SIMILARITY: Belongs to the corticostatin/defensin family.
 DR EMBL; AF191101; AAF04390.1;
 DR EMBL; AF191103; AAF04392.1;

DR EMBL; AF194156; AAF07923.1;
 DR PIR; B59089; B59089.
 DR GO; GO:0050832; P:defense response to fungi; IEA.
 DR GO; GO:0006805; P:xenobiotic metabolism; IEA.
 DR InterPro; IPR002366; Defensin_propep.
 DR Pfam; PF00879; Defensin_propep; 1
 KW Antibiotic; Defensin; Fungicide; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 64 THETA DEFENSIN-1, SUBUNIT B.
 FT CHAIN 65 73 REMOVED IN MATURE FORM.
 FT PROPEP 74 76
 FT DISULFID 68 73
 SQ SEQUENCE 76 AA; 8189 MW; F0B754466156071E CRC64;
 Query Match 53.6%; Score 60; DB 2; Length 76;
 Best Local Similarity 83.3%; Pred. No. 1.4;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 4 RCLCRGVCRCLC 15
 DB 65 RCLCRGVCRCLC 76
 RESULT 4
 Q9D9I2
 ID Q9D9I2 PRELIMINARY; PRT; 168 AA.
 AC Q9D9I2;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
 DE library, clone:1700065I05 product:hypothetical Cysteine-rich region
 DE containing protein, full insert sequence.
 GN Name=4931420D14Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA The FANTOM Consortium,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanishi M.,
RA Okazaki Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK008692; BAB24782.1; -.
DR MGD; MGI:1913992; 4931420D14Rik.
KW Hypothetical protein.
SQ SEQUENCE 168 AA; 18931 MW; 7A2BD279612A5E94 CRC64;
Query Match 53.6%; Score 60; DB 2; Length 168;
Best Local Similarity 56.2%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 3 CRCLCRGVGVCRLCRR 18
DB 66 CRCCCHCRCRCRCRCSR 81
RESULT 5
ID Q6P8T4 PRELIMINARY; PRT; 168 AA.
AC Q6P8T4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schain J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061079; AAH61079.1; -.
KW Hypothetical protein.
SQ SEQUENCE 168 AA; 18947 MW; B00FD3D27B9BF768 CRC64;
Query Match 51.8%; Score 58; DB 2; Length 168;
Best Local Similarity 56.2%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 3 CRCLCRGVGVCRLCRR 18
DB 66 CRCCCHCRCRCRCRCSR 81
RESULT 6
ID O8CH20 PRELIMINARY; PRT; 168 AA.
AC O8CH20;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Basic protein CKT1R3.
GN Name=4931420D14Rik; Synonyms=Ckt1r3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Testis;
RA Xu X., Bai X., Silvius D., Escallier D., McFarland L., Xu P.-X.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF463502; AA015675.1; -.
DR MGD; MGI:1913992; 4931420D14Rik.
SQ SEQUENCE 168 AA; 18957 MW; 8F30D3D27B9BF595 CRC64;
Query Match 51.8%; Score 58; DB 2; Length 168;
Best Local Similarity 56.2%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 3 CRCLCRGVGVCRLCRR 18
DB 66 CRCCCHCRCRCRCRCSR 81
RESULT 7
ID AAH61079 PRELIMINARY; PRT; 168 AA.
AC AAH61079;
DT 14-APR-2004 (TrEMBLrel. 27, Created)
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 14-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schain J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2];
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX Strausberg R.;
 RA Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC061079; AH61079.1; -;
 DR Hypothetical protein.
 SQ SEQUENCE 168 AA; 18947 MW; B00FD3D27B9BF768 CRC64;
 KW Hypothetical protein.
 QY 3 CRLCRRGVCRCLCRR 18
 DB 66 CRCCCYCRCCGCCSR 81
 RESULT 8
 Q9D4K2 PRELIMINARY; PRT; 173 AA.
 AC Q9D4K2;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
 DE library, clone:4931420D14 product:hypothetical cysteine-rich region
 DE containing protein, full insert sequence.
 GN Name:4931420D14Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1];
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=92279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2];
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3];
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4];
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20499374; PubMed=11042159;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5];
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530513; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Kitsuai T., Tashiro H., Itoh M.,
 RA Konno H., Akiyama J., Nishi K., Kizama M., Nishino T., Harada A.,
 RA Sumi N., Ishi Y., Nakamura S., Ikegami T., Kashiwagi K.,
 RA Yanamoto R., Matsumoto H., Sakaguchi S., Ohara E., Watahiki M.,
 RA Fujiwaka S., Inoue K., Izawa M., Koyama T., Matsura S., Kawai J.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RA "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6];
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK016457; BAB30253.1; -;
 DR MGD; MGI:1913992; 4931420D14Rik.
 KW Hypothetical protein.
 SQ SEQUENCE 173 AA; 19581 MW; 147B6F155AC29PDF CRC64;
 QY Query Match 51.8%; Score 58; DB 2; Length 173;
 DB Best Local Similarity 56.2%; Pred. No. 5.2;
 DB Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 DB 3 CRLCRRGVCRCLCRR 18
 DB 66 CRCCCYCRCCGCCSR 81
 RESULT 9
 Q18238 PRELIMINARY; PRT; 188 AA.
 AC Q18238;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein C27A2.5.
 GN ORFNames=C27A2.5;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1];
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Wilson R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2];
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Nhan M.;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U58760; AAK31463.1; --
 DR PIR; T15651; T15651.
 DR HSSP; P10968; 2CWG.
 DR IntAct; Q18238; --
 DR WormPep; C27A2.5; CE04105.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0006952; P:defense response; IEA.
 DR GO; GO:0009613; P:response to pest/pathogen/parasite; IEA.
 DR InterPro; IPR001450; 4Fe4S_ferredoxin.
 DR InterPro; IPR006081; Defensin_alpha.
 DR InterPro; IPR001007; VWF_C.
 DR PROSITE; PS00198; 4FE4S_FERREDOXIN; UNKNOWN_1.
 DR PROSITE; PS00269; DEFENSIN; 1.
 DR PROSITE; PS01208; VWFC_1; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 188 AA; 18878 MW; 0C5D0DC5CA8E0C4B CRC64;
 Query Match 51.8%; Score 58; DB 2; Length 188;
 Best Local Similarity 55.6%; Pred.No. 5.6;
 Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 GVCRCLCRGVGCRCRCR 18
 Db 81 GGGCGCCCRPKCCCCRR 98
 RESULT 10
 Q22048 PRELIMINARY; PRT; 164 AA.
 AC Q22048;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Hypothetical protein T01B7.8.
 GN Name=T01B7.8;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U58854; AAK68158.1; --
 DR HSSP; P10969; 1K7V.
 DR WormPep; C0466.10; CE27649.
 DR InterPro; IPR001450; 4Fe4S_ferredoxin.
 DR InterPro; IPR006081; Defensin_alpha.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR001007; VWF_C.
 DR PROSITE; PS00198; 4FE4S_FERREDOXIN; UNKNOWN_1.
 DR PROSITE; PS00269; DEFENSIN; UNKNOWN_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01208; VWFC_1; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 166 AA; 16971 MW; 9D9D130351BB50F1 CRC64;
 Query Match 50.9%; Score 57; DB 2; Length 166;
 Best Local Similarity 55.6%; Pred.No. 6.7;
 Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 GVCRCLCRGVGCRCRCR 18
 Db 81 GGGCGCCCRPKCCCCRR 98
 RESULT 12
 Q17641 PRELIMINARY; PRT; 197 AA.
 AC Q17641;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Hypothetical protein C04G6.7.
 GN ORFNames=C04G6.7;
 SQ SEQUENCE 164 AA; 16499 MW; C002D48D36C9FCED CRC64;

Query Match 50.9%; Score 57; DB 2; Length 164;
 Best Local Similarity 55.6%; Pred.No. 6.6;
 Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 GVCRCLCRGVGCRCRCR 18
 Db 80 GGGCGCCCRPKCCCCRR 97
 RESULT 11
 Q95QY1 PRELIMINARY; PRT; 166 AA.
 AC Q95QY1;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Hypothetical protein C04G6.10.
 GN ORFNames=C04G6.10;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Wilson R.;
 RL "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Anderson K., Chisoe S.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U58854; AAK68158.1; --
 DR HSSP; P10969; 1K7V.
 DR WormPep; C0466.10; CE27649.
 DR InterPro; IPR001450; 4Fe4S_ferredoxin.
 DR InterPro; IPR006081; Defensin_alpha.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR001007; VWF_C.
 DR PROSITE; PS00198; 4FE4S_FERREDOXIN; UNKNOWN_1.
 DR PROSITE; PS00269; DEFENSIN; UNKNOWN_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01208; VWFC_1; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 166 AA; 16971 MW; 9D9D130351BB50F1 CRC64;
 Query Match 50.9%; Score 57; DB 2; Length 166;
 Best Local Similarity 55.6%; Pred.No. 6.7;
 Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 GVCRCLCRGVGCRCRCR 18
 Db 81 GGGCGCCCRPKCCCCRR 98
 RESULT 12
 Q17641 PRELIMINARY; PRT; 197 AA.
 AC Q17641;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Hypothetical protein C04G6.7.
 GN ORFNames=C04G6.7;
 SQ SEQUENCE 164 AA; 16499 MW; C002D48D36C9FCED CRC64;

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RC MEDLINE=9069613; PubMed=9851916;
 RA Wilson R.;
 RA "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Anderson K., Chisoe S.;
 RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; U55854; AAK68161.1; -.
 DR HSP; P10969; IWGT.
 DR WormPep; C0466.7; CE27652.
 DR InterPro; IPR001450; 4Fe4S ferredoxin.
 DR InterPro; IPR005081; Defensin_alpha.
 DR InterPro; IPR006203; EGF_like.
 DR InterPro; IPR001007; VWF_C.
 DR InterPro; IPR00198; 4Fe4S FERREDOXIN; UNKNOWN_1.
 DR PROSITE; PS00269; DEFENSIN; UNKNOWN_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01208; VWF_1; UNKNOWN_1.
 KW Hypothetical protein.
 KW SEQUENCE 197 AA; 20596 MW; FBSF9457BFB9B8AD CRC64;
 Query Match 50.9%; Score 57; DB 2; Length 197;
 Best Local Similarity 55.6%; Pred. No. 7.8;
 Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 GVCECLCRGVCRCLCR 18
 DB 82 GGCGCCCCCRPRCCCCRR 99

RESULT 13
 Q949G1 PRELIMINARY; PRT; 274 AA.
 AC Q949G1
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein C15ER1PDM.
 GN Name=C15ER1PDM;
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21329048; PubMed=11435398;
 RA Mayer K., Murphy G., Tarchini R., Hambutt R., Volckaert G., Pohl T.,
 RA Dueterhoeft A., Stiekema W., Entian K.D., Terryn N., Lemcke K.,
 RA Haase D., Hall C.R., van Dodeveld A.M., Tingey S.V., Mewes H.W.,
 RA Bevan M., Bancroft I.;
 RA "Conservation of microstructure between a sequenced region of the
 RT genome of rice and multiple segments of the genome of Arabidopsis
 RT thaliana";
 RL Genome Res. 11:1167-1174(2001).
 DR EMBL; AJ307662; CAC39030.1; -.
 DR Gramene; Q949G1; -.
 KW Hypothetical protein.

SQ SEQUENCE 274 AA; 28657 MW; AB547D9BD5470A31 CRC64;
 Query Match 50.9%; Score 57; DB 2; Length 274;
 Best Local Similarity 60.0%; Pred. No. 10;
 Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 RCLCRGVCRCLCR 18
 DB 185 RCCCHRGCCRCRATR 199

RESULT 14
 Q9UI23 PRELIMINARY; PRT; 190 AA.
 AC Q9UI23
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE PRO0529.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Zhang C., Yu Y., Zhang S., Ouyang S., Luo L., Wei H., Zhou G.,
 RA Zhang Y., Liu M., He F.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF111848; AAF16687.1; -.
 SQ SEQUENCE 190 AA; 21480 MW; 4B8104A29AA33844 CRC64;
 Query Match 50.4%; Score 56.5; DB 2; Length 190;
 Best Local Similarity 55.6%; Pred. No. 8.7;
 Matches 10; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

OY 2 VCRCLCRG---VCRCLC 16
 DB 47 VCLCVCVRCVSVCCVCVC 64

RESULT 15
 Q89566 PRELIMINARY; PRT; 413 AA.
 AC Q89566
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE U88
 GN Name=U88;
 OS Human herpesvirus 6.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Roseolovirus.
 OX NCBI_TaxID=10368;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=U1102;
 RX MEDLINE=90112641; PubMed=2153237;
 RA Littler E., Lawrence G., Liu M.Y., Barrell B.G., Arrand J.R.;
 RT "Identification, cloning, and expression of the major capsid protein
 RT gene of human herpesvirus 6";
 RL J. Virol. 64:714-722(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=U1102;
 RX MEDLINE=90080132; PubMed=2152817;
 RA Lawrence G.L., Chee M., Craxton M.A., Gompels U.A., Honess R.W.,
 RA Barrell B.G.;
 RT "Human herpesvirus 6 is closely related to human cytomegalovirus";
 RL J. Virol. 64:287-299(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=U1102;

RX MEDLINE=91226542; PubMed=1851252;
 RA Thomson B.J., Efsthathiou S., Honess R.W.;
 RT "Acquisition of the human adeno-associated virus type-2 rep gene by
 RL human herpesvirus type-6.";
 RN Nature 351:78-80(1991).
 RP [4]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=91333007; PubMed=1651403;
 RA Teo I.A., Griffin B.E., Jones M.D.;
 RT "Characterization of the DNA polymerase gene of human herpesvirus 6.";
 RL J. Virol. 65:4670-4680(1991).
 RN [5]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=91374590; PubMed=1654446;
 RA Martin M.E.D., Nicholas J., Thomson B.J., Newman C., Honess R.W.;
 RT "Identification of a transactivating function mapping to the putative
 RL immediate-early locus of human herpesvirus 6.";
 RN J. Virol. 65:5381-5390(1991).
 RP [6]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=91237802; PubMed=1851860;
 RA Chang C.K., Balachandran N.;
 RT "Identification, characterization, and sequence analysis of a cDNA
 RL encoding a phosphoprotein of human herpesvirus 6.";
 RN J. Virol. 65:2884-2894(1991).
 RP [7]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=92148942; PubMed=1310766;
 RA Geng Y., Chandran B., Josephs S.F., Wood C.;
 RT "Identification and characterization of a human herpesvirus 6 gene
 RL segment that trans activates the human immunodeficiency virus type 1
 RT promoter.";
 RN J. Virol. 66:1564-1570(1992).
 RP [8]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=92260671; PubMed=1374813;
 RA Neipel F., Ellinger K., Fleckenstein B.;
 RT "Gene for the major antigenic structural protein (p100) of human
 RL herpesvirus 6.";
 RN J. Virol. 66:3918-3924(1992).
 RP [9]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=92333249; PubMed=1321206;
 RA Efsthathiou S., Lawrence G.L., Brown C.M., Barrell B.G.;
 RT "Identification of homologs to the human cytomegalovirus US22 gene
 RL family in human herpesvirus-6.";
 RN J. Gen. Virol. 73:1661-1671(1992).
 RP [10]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=93187613; PubMed=8383182;
 RA Ellinger K., Neipel F., Foa-Romasi L., Campadelli-Fiume G.,
 RT Fleckenstein B.;
 RL "The glycoprotein B homologue of human herpesvirus 6.";
 RN J. Gen. Virol. 74:495-500(1993).
 RP [11]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=92333248; PubMed=1321205;
 RA Thomson B.J., Honess R.W.;
 RT "The right end of the unique region of the genome of human herpesvirus
 RL 6 U1102 contains a candidate immediate early gene enhancer and a
 RT homologue of the human cytomegalovirus US22 gene family.";
 RN J. Gen. Virol. 73:1649-1660(1992).
 RP [12]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=94167865; PubMed=8123364;
 RA
 RT
 RL
 RN
 RP
 RX
 MEDLINE=93091236; PubMed=1333836;
 RA Gompels U.A., Carss A.L., Sun N., Arrand J.R.;
 RT "Infectivity determinants encoded in a conserved gene block of human
 RL herpesvirus-6.";
 RN DNA Seq. 3:25-39(1992).
 RP [13]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=94181269; PubMed=8134119;
 RA Thompson J., Choudhury S., Kashanchi F., Doniger J., Berneman Z.,
 RT Frenkel N., Rosenthal L.J.;
 RL "A transforming fragment within the direct repeat region of human
 RL Oncogene 9:1167-1175(1994).
 RP [14]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=93224892; PubMed=8385692;
 RA Gompels U.A., Carrigan D.R., Carss A.L., Arno J.;
 RT "Two groups of human herpesvirus 6 identified by sequence analyses of
 RL laboratory strains and variants from Hodgkin's lymphoma and bone
 RL marrow transplant patients.";
 RN J. Gen. Virol. 74:613-622(1993).
 RP [15]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=93323202; PubMed=7687301;
 RA Pfeiffer B., Berneman Z.N., Neipel F., Chang C.K., Tirwatnang S.,
 RT Chandran B.;
 RL "Identification and mapping of the gene encoding the glycoprotein
 RL complex gp82-gp105 of human herpesvirus 6 and mapping of the
 RL neutralizing epitope recognized by monoclonal antibodies.";
 RN J. Virol. 67:4611-4620(1993).
 RP [16]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=93337170; PubMed=7687803;
 RA Pellett P.E., Sanchez-Martinez D., Dominguez G., Black J.B., Anton E.,
 RT Greenamoyer C., Dambaugh T.R.;
 RL "A strongly immunoreactive virion protein of human herpesvirus 6
 RT variant B strain Z29: identification and characterization of the gene
 RL and mapping of a variant-specific monoclonal antibody reactive
 RL epitope.";
 RN Virology 195:521-531(1993).
 RP [17]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=94025598; PubMed=8212582;
 RA Jones M., Teo I.;
 RT "Identification and analysis of the transport/capsid assembly protein
 RL (tp/cap) gene of human herpesvirus-6 (HHV-6).";
 RN Virology 197:449-454(1993).
 RP [18]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=93389439; PubMed=8397282;
 RA Liu D.X., Gompels U.A., Nicholas J., Lellicott C.;
 RT "Identification and expression of the human herpesvirus 6 glycoprotein
 RL H and interaction with an accessory 40K Glycoprotein.";
 RN J. Gen. Virol. 74:1847-1857(1993).
 RP [19]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=94118404; PubMed=8289364;
 RA Nicholas J., Martin M.;
 RT "Nucleotide sequence analysis of a 38.5-kilobase-pair region of the
 RL genome of human herpesvirus 6 encoding human cytomegalovirus
 RL immediate-early gene homologs and transactivating functions.";
 RN J. Virol. 68:597-610(1994).
 RP [20]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN=U1102;
 RX MEDLINE=94167865; PubMed=8123364;
 RA

RA Zhou Y., Chang C.K., Qian G., Chandran B., Wood C.;
 RT "trans-activation of the HIV promoter by a cDNA and its genomic clones
 RL of human herpesvirus-6.";
 RL Virology 199:311-322(1994).
 [21]
 RN SEQUENCE FROM N.A.
 RP STRAIN=U1102;
 RC MEDLINE=95146989; PubMed=7844567;
 RX Gompels U.A., Macaulay H.A.;
 RA "Characterisation of human telomeric repeat sequences from human
 RT herpesvirus-6 and relationship to replication.";
 RL J. Gen. Virol. 76:451-458(1995).
 [22]
 RN SEQUENCE FROM N.A.
 RP STRAIN=U1102;
 RC MEDLINE=94202288; PubMed=8151770;
 RX Thomson B.J., Dewhurst S., Gray D.;
 RA "Structure and heterogeneity of the a sequences of human herpesvirus 6
 RT strain variants U1102 and Z29 and identification of human telomeric
 RL repeat sequences at the genomic termini.";
 RL J. Virol. 68:3007-3014(1994).
 [23]
 RN SEQUENCE FROM N.A.
 RP STRAIN=U1102;
 RC

Query Match 50.4%; Score 56.5; DB 2; Length 413;
 Best Local Similarity 66.7%; Pred. No. 17;
 Matches 10; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

Qy 2 VCRLCRRGVCRCLC 16
 |||||
 Db 123 VCACLC---VCACLC 134

RESULT 16

Q6RY99 ID Q6RY99 PRELIMINARY; PRT; 512 AA.
 AC Q6RY99;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Gamma-hydroxybutyrate receptor.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Hippocampus;
 RA Andriampandry C., Taleb O., Viry S., Muller C., Humbert J.P.,
 RA Gobaille S., Aunis D., Maitre M.;
 RA "Cloning and characterization of a rat brain receptor that binds the
 RT endogenous neuromodulator gamma-hydroxybutyrate.";
 RL FASEB J. 0:0-0(2004).
 DR ENBL; AY485933; AAR24072.1; -;
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR001209; Ribosomal_S14.
 DR InterPro; IPR008952; Tetraspanin.
 DR Pfam; PF00335; Tetraspanin; I.
 DR PRINTS; PR00259; TWFOUR.
 DR PROSITE; PS00527; RIBOSOMAL_S14; UNKNOWN_1.
 KW Receptor.
 SQ SEQUENCE 512 AA; 56146 MW; 44DD12B937C40D60 CRC64;

Query Match 49.6%; Score 55.5; DB 2; Length 512;
 Best Local Similarity 64.7%; Pred. No. 28;
 Matches 11; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 2 VCRLCRRGVCRCLC 18
 |||||
 Db 385 VCVCVCVR-VCLCLCVR 400

RESULT 17

AAR24072 ID AAR24072 PRELIMINARY; PRT; 512 AA.
 AC AAR24072;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Gamma-hydroxybutyrate receptor.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Hippocampus;
 RA Andriampandry C., Taleb O., Viry S., Muller C., Humbert J.P.,
 RA Gobaille S., Aunis D., Maitre M.;
 RA "Cloning and characterization of a rat brain receptor that binds the
 RT endogenous neuromodulator gamma-hydroxybutyrate.";
 RL FASEB J. 0:0-0(2004).
 DR ENBL; AY485933; AAR24072.1; -;
 DR KW Receptor.
 SQ SEQUENCE 512 AA; 56146 MW; 44DD12B937C40D60 CRC64;

Query Match 49.6%; Score 55.5; DB 2; Length 512;
 Best Local Similarity 64.7%; Pred. No. 28;
 Matches 11; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 2 VCRLCRRGVCRCLC 18
 |||||
 Db 385 VCVCVCVR-VCLCLCVR 400

RESULT 18

Q6ZQS2 ID Q6ZQS2 PRELIMINARY; PRT; 201 AA.
 AC Q6ZQS2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein FLJ45585.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thalamus;
 RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
 RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
 RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
 RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR ENBL; AK128797; BAC87611.1; -;
 DR InterPro; IPR006209; EGF-like.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 SQ SEQUENCE 201 AA; 20938 MW; 3408F8E817AA0500 CRC64;

Query Match 49.1%; Score 55; DB 2; Length 201;
 Best Local Similarity 53.3%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 VCRLCRRGVCRCLC 16
 |||||
 Db 58 VCLCVCVSVCLCVC 72

RESULT 19

BAC87611

| | | | | |
|--------|--|--|---------|---------|
| ID | BAC87611 | PRELIMINARY; | PRT; | 201 AA. |
| AC | BAC87611; | | | |
| DT | 02-MAR-2004 | (TREMBLrel. 27, Created) | | |
| DE | 02-MAR-2004 | (TREMBLrel. 27, Last sequence update) | | |
| DT | 02-MAR-2004 | (TREMBLrel. 27, Last annotation update) | | |
| DE | CDNA FLJ45585 fis. clone BRTHA3013882. | | | |
| OS | Homo sapiens (Human) | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Thalamus; | | | |
| RA | Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S., | | | |
| RA | Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M., | | | |
| RA | Sugiyama T., Irie R., Osuki T., Sato H., Wakamatsu A., Ishii S., | | | |
| RA | Yamanoto J., Kawai-Hio Y., Saito K., Nishikawa T., | | | |
| RA | Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., | | | |
| RA | Wagatsuna M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., | | | |
| RA | Suzuki Y., Sugano S., Nagahari K., Masubo Y., Nagai K., Isogai T., | | | |
| RT | "NEED human cDNA sequencing project"; | | | |
| RL | Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; AKI28797; BAC87611.1; -- | | | |
| SQ | SEQUENCE 201 AA; 20938 MW; 3408FBE817AA0500 CRC64; | | | |
| | Query Match | 49.1%; Score 55; DB 2; Length 201; | | |
| | Best Local Similarity | 53.3%; Pred. No. 14; | | |
| | Matches | 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0; | | |
| QY | 2 VRCCLRRGVCRCLC 16 | | | |
| Dd | ::: :: | | | |
| Db | 58 VCLCVCLVSVCLVC 72 | | | |
| | RESULT 20 | | | |
| Q7PDW6 | PRELIMINARY; | PRT; | 212 AA. | |
| ID | Q7PDW6; | | | |
| AC | Q7PDW6; | | | |
| DT | 01-MAR-2004 | (TREMBLrel. 26, Created) | | |
| DT | 01-MAR-2004 | (TREMBLrel. 26, Last sequence update) | | |
| DT | 01-MAR-2004 | (TREMBLrel. 26, Last annotation update) | | |
| DE | ENSANGP0000024925 (Fragment). | | | |
| GN | Name=ENSANGP0000020976; | | | |
| OS | Anopheles gambiae str. PEST. | | | |
| OC | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | | | |
| OC | Neoptera; Endopterygota; Diptera; Nematocera; Anophelinae; | | | |
| OC | NCBI_TaxID=180454; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=PEST; | | | |
| RA | Anopheles Genome Sequencing Consortium; | | | |
| RL | Submitted (APR-2003) to the EMBL/GenBank/DBJ databases. | | | |
| CC | -! CAUTION: The sequence shown here is derived from an | | | |
| CC | EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is | | | |
| CC | preliminary data. | | | |
| DR | EMBL; AAA01000199; EAA45965.1; -- | | | |
| FT | NON_TER | 212 | | |
| SQ | SEQUENCE 212 AA; 23844 MW; 28984CAECB21D496 CRC64; | | | |
| | Query Match | 48.7%; Score 54.5; DB 2; Length 212; | | |
| | Best Local Similarity | 60.0%; Pred. No. 17; | | |
| | Matches | 9; Conservative 1; Mismatches 2; Indels 3; Gaps 1; | | |
| QY | 3 CRCLRRGVCRCLCR 17 | | | |
| Dd | : | | | |
| Db | 33 CRRCQC---CRRCRC 44 | | | |
| | RESULT 21 | | | |
| Q7HWV7 | PRELIMINARY; | PRT; | 212 AA. | |
| ID | Q7HWV7; | | | |
| AC | Q7HWV7; | | | |
| DT | 01-OCT-2003 | (TREMBLrel. 25, Created) | | |

| | | | | |
|----|--|--|--|--|
| DT | 01-OCT-2003 | (TREMBLrel. 25, Last sequence update) | | |
| DT | 01-MAR-2004 | (TREMBLrel. 26, Last annotation update) | | |
| DE | Hypothetical protein R13H4.8; | | | |
| OS | Name=R13H4.8; | | | |
| GN | Caenorhabditis elegans. | | | |
| OC | Eukaryota; Metazoa; Nematoda; Chromadorea; Rhaditida; Rhaditoidea; | | | |
| OC | Rhaditidae; Peloderinae; Caenorhabditis. | | | |
| OX | NCBI_TaxID=6239; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=Bristol N2; | | | |
| RX | MEDLINE=99069613; PubMed=9851916; | | | |
| RA | none; | | | |
| RT | "Genome sequence of the nematode C.elegans: A platform for | | | |
| RT | investigating biology."; | | | |
| RL | Science 282:2012-2018(1998). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=Bristol N2; | | | |
| RA | Kershaw J.K.; | | | |
| RL | Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; Z81579; CAB17915.1; -- | | | |
| DR | InterPro; IPR001450; 4Fe4S_ferredoxin. | | | |
| DR | InterPro; IPR006209; EGF-like. | | | |
| DR | InterPro; IPR001007; VMF_C. | | | |
| DR | PROSITE; PS00198; 4FE4S_FERREDOXIN; UNKNOWN_1. | | | |
| DR | PROSITE; PS00022; EGF_1; UNKNOWN_1. | | | |
| DR | PROSITE; PS01208; VMFC_1; UNKNOWN_1. | | | |
| XW | Hypothetical protein. | | | |
| SQ | SEQUENCE 212 AA; 22857 MW; 22F62CE4073CCA4E CRC64; | | | |
| | Query Match | 48.2%; Score 54; DB 2; Length 212; | | |
| | Best Local Similarity | 56.2%; Pred. No. 20; | | |
| | Matches | 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0; | | |

DB 728 GFAPCTCRFPVASCFCR 744

RESULT 23

Q8IMJ2 PRELIMINARY; PRT; 905 AA.

AC Q8IMJ2

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE CG31029-PA.

GN ORFNames=CG31029;

OS Drosophila melanogaster (fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophiliidae; Drosophila.

OX NCBI_TaxID=7227;

RT "The genome sequence of Drosophila melanogaster."

RL NCBI_TaxID=7227;

RP SEQUENCE FROM N.A.

RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle K., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,

RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Moberg C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodager, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,

RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).

RP SEQUENCE FROM N.A.

RA MEDLINE=2426065; PubMed=12537568;

RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,

RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,

RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,

RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,

RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,

RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;

RT "Finishing a whole-genome shotgun: release 3 of the Drosophila melanogaster euchromatic genome sequence."

RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

RP SEQUENCE FROM N.A.

RA MEDLINE=22426070; PubMed=12537573;

RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J., Svirskaas R.,

RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,

RA Ashburner M., Celniker S.E.;

RT "The transposable elements of the Drosophila melanogaster euchromatic genome perspective."

RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

RP SEQUENCE FROM N.A.

RA MEDLINE=22426069; PubMed=12537572;

RA Mira S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,

RA Smith C.D., Rupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,

RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,

RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,

RA Lewis S.E.;

RT "Annotation of the Drosophila melanogaster euchromatic genome: a systematic review."

RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).

RP SEQUENCE FROM N.A.

RA FLYBASE;

RL Submitted (SSP-2002) to the EMBL/GenBank/DBJ databases.

RN [6]

RP SEQUENCE FROM N.A.

RG FLYBASE;

RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

RL EMBL; AE003772; AAN14217.1; -.

DR IntAct; Q8IMJ2; -.

DR FLYBASE; FBgn0051029; CG31029.

DR SEQUENCE 905 AA; 102461 MW; 59471B320E041C20 CRC64;

SQ

Query Match 48.2%; Score 54; DB 2; Length 905;

Best Local Similarity 52.9%; Pred. NO. 70;

Matches 9; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GVCRLCGRGVCRLCR 17

DB 728 GFAPCTCRFPVASCFCR 744

RESULT 24

Q7SXV0 PRELIMINARY; PRT; 379 AA.

AC Q7SXV0

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Hypothetical protein zgc:63759.

GN Name=zgc:63759;

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI_TaxID=7955;

RP SEQUENCE FROM N.A.

RC STRAIN=AB; TISSUE=Whole body;

RA MEDLINE=22389257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Whole body;
 RA Strausberg R.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC055236; AAH55236.1; -
 DR InterPro: IPR002172; LDL_receptor_A.
 DR Pfam: PF00057; Ldl_recept_a; 6.
 DR SMART: SM00192; LDLA_1; 6.
 DR PROSITE: PS01209; LDLRA_1; 6.
 DR PROSITE: PS00068; LDLRA_2; 6.
 KW Hypothetical protein.
 SQ SEQUENCE 379 AA; 40591 MW; D763CB41A9F752C6 CRC64;

Query Match 47.8%; Score 53.5; DB 2; Length 379;
 Best Local Similarity 45.5%; Pred. No. 38;
 Matches 10; Conservative 2; Mismatches 3; Indels 7; Gaps 1;

QY 2 VCRCL-----CRGVGRCCLC 16
 DB 345 VCLCVTEVLSDPCRCGVGVVC 366

RESULT 25
 O57352 PRELIMINARY; PRT; 418 AA.
 ID O57352
 AC O57352
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Vascular endothelial growth factor C precursor.
 GN Name=VEGF-C;
 OS Coturnix coturnix (Common quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Coturnix.

NCBI_TaxID=9091;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98167900; PubMed=9435294;
 RA Richmann A., Corbel C., Jaffredo T., Breant V., Joukov V., Kumar V.,
 RA Alitalo K., Le Douarin N.M.;
 RA "Avian VEGF-C: cloning, embryonic expression pattern and stimulation
 RT of the differentiation of VEGFR2 expressing endothelial cell
 RT precursors";
 RL Development 125:743-752(1998).
 CC 1- SIMILARITY: Belongs to the PDGF/VEGF growth factor family.
 DR EMBL: Y15837; CAAV5799.1; -.
 DR HSSP: P49763; 1FZV.

DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0008083; F:growth factor activity; IEA.
 DR GO: GO:0008151; P:cell growth and/or maintenance; IEA.
 DR InterPro: IPR004153; CXCXC repeat.
 DR InterPro: IPR002400; GF_cysknor.
 DR Pfam: PF03128; CXCXC; 6.
 DR Pfam: PF00341; PDGF; 1.
 DR PRINTS: PR00438; GFCYXKNOT.
 DR ProDom: PD001629; PD_growth_factor; 1.
 DR SMART: SM00141; PDGF_1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS00278; PDGF_2; 1.
 KW Growth factor; Mitogen; Signal.
 FT SIGNAL 1 31 Potential.
 FT CHAIN 111 418 vascular endothelial growth factor C.
 SQ SEQUENCE 418 AA; 46839 MW; 099BFCC79151BF2B CRC64;

Query Match 47.8%; Score 53.5; DB 2; Length 418;
 Best Local Similarity 26.7%; Pred. No. 41;
 Matches 8; Conservative 6; Mismatches 1; Indels 15; Gaps 1;
 QY 3 CRCLCRGV-----CRCLCR 17
 DB 290 CQCCKGVPRPISCGPHKELDRASCCQCK 319

RESULT 26

VEGC_HUMAN
 ID_VEGC_HUMAN STANDARD; PRT; 419 AA.
 AC P49767;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Vascular endothelial growth factor C precursor (VEGF-C) (Vascular
 DE endothelial growth factor related protein) (VRP) (Flt4 ligand) (Flt4-
 DE L).
 GN Name=VEGFC;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 103-120.
 RX MEDLINE=96178224; PubMed=8617204;
 RA Joukov V., Pajusola K., Kaipainen A., Chilov D., Lahtinen I., Kukk E.,
 RA Saksela O., Kalkkinen N., Alitalo K.;
 RT "A novel vascular endothelial growth factor, VEGF-C, is a ligand for
 RT the Flt4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases";
 RL EMBO J. 15:290-298(1996).
 RN [2]
 RP ERZATUM.
 RX MEDLINE=96203094; PubMed=8612600;
 RA Joukov V., Pajusola K., Kaipainen A., Chilov D., Lahtinen I., Kukk E.,
 RA Saksela O., Kalkkinen N., Alitalo K.;
 RL EMBO J. 15:1751-1751(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Glial tumor;
 RX MEDLINE=96312526; PubMed=8700872;
 RA Lee J., Gray A., Yuan J., Luoh S.-M., Avraham H., Wood W.I.;
 RT "Vascular endothelial growth factor-related protein: a ligand and
 RT specific activator of the tyrosine kinase receptor Flt4";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97388482; PubMed=9247316;
 RA Wang J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,
 RA Giannotti J., Finnerty H., Zoller R., Beier D.R., Leak L.V.,
 RA Turner K.J., Wood C.R.;
 RT "Characterization of murine Flt4 ligand/VEGF-C";
 RL Oncogene 15:613-618(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=97388482; PubMed=9247316;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny K.C., Hale S., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP SEQUENCE OF 32-41; 112-121 AND 228-233, AND MUTAGENESIS OF ARG-227.
RX MEDLINE=97377029; Pubmed=9233800;
RA Joukov V., Sorsa T., Kumar V., Jellech M., Claesson-Welsh L., Cao Y.,
RA Saksela O., Kalkkinen N., Alitalo K.;
RT "Proteolytic processing regulates receptor specificity and activity of
RT VEGF-C";
RL EMBO J. 16:3898-3911(1997).
RN [7]
RP SEQUENCE OF 32-46.
RA Zhang Z., Henzel W.;
RT "Signal peptide prediction based on analysis of experimentally
RT verified cleavage sites.";
RL Submitted (JUN-2004) to Swiss-Prot.
CC -1- FUNCTION: Growth factor active in angiogenesis, and endothelial
CC cell growth, stimulating their proliferation and migration and
CC also has effects on the permeability of blood vessels. May
CC function in angiogenesis of the venous and lymphatic vascular
CC systems during embryogenesis, and also in the maintenance of
CC differentiated lymphatic endothelium in adults. Binds and
CC activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.
CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Spleen, lymph node, thymus, appendix, bone
CC marrow, heart, placenta, ovary, skeletal muscle, prostate, testis,
CC colon and small intestine and fetal liver, lung and kidney, but
CC not in peripheral blood lymphocyte.
CC -1- PTM: Undergoes a complex proteolytic maturation which generates a
CC variety of processed secreted forms with increased activity toward
CC VEGFR-3, but only the fully processed form could activate VEGFR-2.
CC VEGF-C first form an antiparallel homodimer linked by disulfide
CC bonds. Before secretion, a cleavage occurs between arg-227 and
CC ser-228 producing an heterotetramer. The next extracellular step
CC of the processing removes the N-terminal propeptide. Finally the
CC mature VEGF-C is composed mostly of two VEGF homology domains
CC (VHDs) bound by non-covalent interactions.
CC -1- SIMILARITY: Belongs to the PDGF/VEGF growth factor family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X94216; CAA63907.1; -;
DR EMBL; U43142; AAB65214.1; -;
DR EMBL; U58111; AAB02909.1; -;
DR EMBL; BC035212; AAB35212.1; -;
DR PIR; S69207; S69207.
DR HSSP; P49763; 1PZV.
DR Genew; HGNC:12682; VEGFC.
DR MTM; 601528.
DR GO; GO:0007515; P:lymph gland development; TAS.
DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR GO; GO:0006929; P:substrate-bound cell migration; TAS.
DR InterPro; IPR004153; CXKC repeat.
DR InterPro; IPR002400; GF cystknot.
DR InterPro; IPR000072; PD_growth_factor.
DR Pfam; PF03128; CXKC; 4.
DR Pfam; PF00341; PDGF; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR ProDom; PD001629; PD_growth_factor; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS00278; PDGF_2; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS00278; PDGF_2; 1.

KW Angiogenesis; Cleavage on pair of basic residues;
KW Direct protein sequencing; Glycoprotein; Growth factor; Mitogen;
KW Multigene family; Repeat; Signal.
FT SIGNAL 1 31
FT PROPEP 32 111 Or 102.
FT CHAIN 112 227 Vascular endothelial growth factor C.
FT PROPEP 228 419
FT DOMAIN 280 362
FT REPEAT 280 295
FT REPEAT 304 319
FT REPEAT 328 343
FT REPEAT 347 362
FT DISULFID 131 173 By similarity.
FT DISULFID 162 209 By similarity.
FT DISULFID 166 211 Interchain (By similarity).
FT DISULFID 165 156 Interchain (By similarity).
FT CARBOHYD 175 175 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 205 205 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 240 240 N-linked (GlcNAc...) (Potential).
FT MUTAGEN 227 227 R-S: No proteolytic processing and lower effect on VEGFR-2 and VEGFR-3.
SQ SEQUENCE 419 AA; 46883 MW; 9F598719DB3E014F CRC64;
Query Match 47.8%; Score 53.5; DB 1; Length 419;
Best Local Similarity 26.7%; Pred. No. 41;
Matches 8; Conservative 6; Mismatches 1; Indels 15; Gaps 1;
QY 3 CRCLCRGV-----CRCLCR 17
DB 291 CQCVCRLAPASCGPHKELDRNSCCQCK 320
RESULT 27
Q6FH59 PRELIMINARY; PRT; 419 AA.
AC Q6FH59;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE VEGFC protein.
GN Name=VEGFC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Halleck A., Ebert L., Mkundinya M., Schick M., Eisenstein S.,
RA Neubert P., Kstrang K., Schatten R., Shen B., Henze S., Mar W.,
RA Korn B., Zuo D., Hu Y., LaBaer J.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the PDGF/VEGF growth factor family.
DR EMBL; CR541897; CAC46695.1; -;
DR InterPro; IPR004153; CXKC repeat.
DR InterPro; IPR002400; GF cystknot.
DR InterPro; IPR000072; PD_growth_factor.
DR Pfam; PF03128; CXKC; 4.
DR Pfam; PF00341; PDGF; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR ProDom; PD001629; PD_growth_factor; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS00278; PDGF_2; 1.
KW Growth factor; Mitogen.
SQ SEQUENCE 419 AA; 46943 MW; 9F598703C13E1B55 CRC64;
Query Match 47.8%; Score 53.5; DB 2; Length 419;
Best Local Similarity 26.7%; Pred. No. 41;
Matches 8; Conservative 6; Mismatches 1; Indels 15; Gaps 1;
QY 3 CRCLCRGV-----CRCLCR 17

| | |
|--|--|
| RA | Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., |
| RA | Sekine M., Obayashi M., Nishi T., Shibehara T., Tanaka T., Ishii S., |
| RA | Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K., |
| RA | Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., |
| RA | Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M., |
| RA | Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., |
| RA | Abe K., Kanihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., |
| RA | Ninomiya K., Ishibaeshi T., Yanashita H., Murakawa K., Fujimori K., |
| RA | Tanai H., Kimiguchi S., Watanabe M., Hiraoka S., Chiba Y., Ishida S., |
| RA | Ono Y., Takiguchi S., Watanabe S., Yosida M., Houta T., Kusano J., |
| RA | Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y., |
| RA | Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., |
| RA | Mushahino Y., Yuuki H., Oshima A., Sasaki N., Aotsuka S., |
| RA | Yoshikawa K., Matsumawa H., Ichihara T., Shiohata N., Sano S., |
| RA | Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O., |
| RA | Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., |
| RA | Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., |
| RA | Fujimori Y., Komiya M., Tashiro H., Tanigami A., Fujiwara T., |
| RA | Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., |
| RA | Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., |
| RA | Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., |
| RA | Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., |
| RA | Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., |
| RA | Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., |
| RA | Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., |
| RA | Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S., |
| RT | "complete sequencing and characterization of 21,243 full-length human |
| RT | cdnas"; |
| RL | Nat. Genet. 36:40-45(2004). |
| DR | EMBL: AK022609; EBI4128.1; -- |
| DR | SEQUENCE 307 AA; 32780 MW; 4CC18ACD39B03AC3 CRC64; |
| Query Match 47.3%; Score 53; DB 2; Length 307; | |
| Best Local Similarity 57.9%; Pred. No. 36; | |
| Matches 11; Conservative 0; Mismatches 4; Indels 4; Gaps | |
| Qy | 2 VRCICRG-----VCRCLC 16 |
| | |
| Db | 116 VCGCLCVGAHLCVVCVLC 134 |
| RESULT 30 | |
| ID | Q3W0A0 PRELIMINARY; PRT; 881 AA. |
| QD | Q3W0A0; |
| AC | Q3W0A0; |
| DT | 01-MAY-2000 (TrEMBLrel. 13, Created) |
| DT | 01-JUN-2003 (TrEMBLrel. 24, Last sequence update) |
| DT | 01-MAR-2004 (TrEMBLrel. 26, Last annotation update) |
| DE | CS2086-PB. |
| GN | Names-drp1; ORFNames=CG2086; |
| OS | Drosophila melanogaster (Fruit fly). |
| OC | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; |
| OC | Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; |
| OC | Ephydroidea; Drosophilidae; Drosophila. |
| OC | NCBI_TaxID=7227; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE=20196006; PubMed=10731132; |
| RA | Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D., |
| RA | Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., |
| RA | George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., |
| RA | Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X., |
| RA | Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D., |
| RA | Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L., |
| RA | Abriel J.F., Agbayani A., An H.J., Andrews-Frannkoch C., Baldwin D., |
| RA | Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M., |
| RA | Beeson K.Y., Benos P.V., Berwan B.P., Bhandari D., Bolshakov S., |
| RA | Borkova D., Bochan M.R., Bouck J., Brokstein P., Brottier P., |
| RA | Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I., |
| RA | Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., |
| RA | de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., |
| RA | Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov S.C., Dunn P., |

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.J., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodagat, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "the genome sequence of *Drosophila melanogaster*."
 RT Science 287:2185-2195(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=22426065; PubMed=12537568;
 RX Celinker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RA "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
 RT melanogaster euchromatic genome sequence."
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=22426070; PubMed=12537573;
 RX Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celinker S.E.;
 RT "the transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomic perspective."
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=22426069; PubMed=12537572;
 RX Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celinker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review."
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RN SEQUENCE FROM N.A.
 RP FLYBASE;
 RG Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE FROM N.A.
 RP FLYBASE;
 RG Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AEO03472; AAF47553.2; -;
 DR HSP; Q12780; IHA6.
 DR FlyBase; FBgn0027594; dprp.
 DR GO; GO:0008345; P:larval locomotory behavior; IMP.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR002049; Laminin_EGF.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00053; Laminin_EGF; 6.
 DR PRINTS; PR00011; EGFLAMININ.
 DR SMART; SM00180; EGF_Lam; 3.
 DR PROSITE; PS00022; EGF_1; 11.
 DR PROSITE; PS01186; EGF_2; 13.
 DR PROSITE; PS00026; EGF_3; 7.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 KW Laminin EGF-like domain.
 SQ SEQUENCE 881 AA; 96380 MW; 52196D164F52F5C1 CRC64;
 Query Match 47.3%; Score 53; DB 2; Length 881;
 Best Local Similarity 64.3%; Pred. No. 91;
 Matches 9; Conservative 3; Mismatches 0; Indels 2; Gaps 1;
 QY 4 RCLRR--GVCRCL 15
 DB 409 RCVCCKQDWGVCRL 422
 RESULT 31
 Q7PRP5 PRELIMINARY; PRT; 1823 AA.
 ID AC Q7PRP5;
 DT 01-MAR-2004 (TREMBlrel. 26, Created)
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
 DE ENSANGP0000019046 (Fragment).
 GN Name=ENSANGP0000019046 (Fragment).
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.
 OX NCBI_TaxID=180454;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PEST;
 RA Anopheles Genome Sequencing Consortium;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAB0100847; EAA06854.2; -;
 DR InterPro; IPR000152; Asx_Hydroxyl_S.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR003410; Hyalin.
 DR InterPro; IPR001791; Laminin_G.
 DR InterPro; IPR001759; Pentaxin.
 DR InterPro; IPR001211; PhospholipaseA2.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF00008; EGF; 5.
 DR Pfam; PF02494; HYR; 2.
 DR Pfam; PF00354; Pentaxin; 1.
 DR Pfam; PF00084; Sushi; 8.
 DR Pfam; PF00092; VWA; 1.
 DR PRINTS; PR00895; PENTAXIN.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR PRODOM; PD002153; Pentaxin; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_5.
 DR PROSITE; PS01186; EGF_2; 4.
 DR PROSITE; PS50025; EGF_3; 5.
 DR PROSITE; PS50825; HYR_2.
 DR PROSITE; PS50025; LAM_G_DOMAIN; 1.
 DR PROSITE; PS00118; PA2_HIS; UNKNOWN_1.
 DR PROSITE; PS50923; SUSHI; 8.
 DR PROSITE; PS50234; VWF_A; 1.
 FT NON_TER 1
 FT NON_TER 1823
 SQ SEQUENCE 1823 AA; 200582 MW; 5740C4C700804379 CRC64;

```

Query Match      47.3%; Score 53; DB 2; Length 1823;
Best Local Similarity 56.2%; Pred. No. 1.7e+02;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      2 VGCRLCRGVCRCLC 17
DB      262 VCSCLTTSCHVRCICR 277

RESULT 32
QYU02      PRELIMINARY; PRT; 101 AA.
AC QYU02;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=TB927.2.4940;
OS Trypanosoma brucei.
OC Eukaryota; Eulenzozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22789168; PubMed=12907728;
RA El-Sayed N.M.A., Ghedin E., Song J., MacLeod A., Bringaud F.,
RA Larkin C., Wailes D., Peterson J., Hou L., Taylor S., Tweedie A.,
RA Bibeau N., Khalaf H.G., Lin X., Mason T., Hannick L., Catter B.,
RA Blandin G., Bartholomeu D., Simpson A.J., Kaul S., Zhao H., Pal G.,
RA Van Aken S., Utterback T., Haas B., Koo H.L., Umayam L., Suh B.,
RA Gerard C., Leech V., Qi R., Zhou S., Schwartz D., Feldblyum T.,
RA Salzberg S., Tait A., Turner M.R., Ullu E., White O., Melville S.,
RA Adams M.D., Fraser C.M., Donelson J.E.;
RT "The sequence and analysis of Trypanosoma brucei chromosome II.";
RL Nucleic Acids Res. 31:4856-4863 (2003).
DR EMBL; AB017169; AAQ15952.1; -
KW Hypothetical protein.
SQ SEQUENCE 101 AA; 10880 MW; E63D8E62A9CA6748 CRC64;

Query Match      46.9%; Score 52.5; DB 2; Length 101;
Best Local Similarity 56.2%; Pred. No. 16;
Matches 9; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY      1 GVCRLCRGVCRCLC 16
DB      4 GVCVCVC---VCVCVC 16

RESULT 33
Q8C9N2      PRELIMINARY; PRT; 118 AA.
AC Q8C9N2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched
DE library, clone:A630032M05 product:hypothetical protein, full insert
DE sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.,
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RN RIKEN FANTOM Consortium;

```

```

RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Thymus;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Thymus;
RX Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ono N., Sano H.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tegami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK041730; EAC31047.1; -
DR InterPro; IPR002453; Beta_cubulin.
DR PROSITE; PS00228; TUBULIN_B_AUTOREG; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 118 AA; 13452 MW; CA98E1C6B01CD77F CRC64;

```

```

Query Match      46.9%; Score 52.5; DB 2; Length 118;
Best Local Similarity 56.2%; Pred. No. 18;
Matches 9; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY      1 GVCRLCRGVCRCLC 16
DB      51 GVCVCVC---VCVCVC 63

RESULT 34
Q6R5G9      PRELIMINARY; PRT; 133 AA.
AC Q6R5G9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Sharov A.A., Piao Y., Matoba R., Dudekula D.B., Qian Y., VanBuren V.,
 RA Falco G., Martin P.R., Stagg C.A., Bassey U.C., Wang Y., Carter M.G.,
 RA Hamatani T., Aiba K., Akutsu H., Sharova L., Tanaka T.S., Kimber W.L.,
 RA Yoshikawa T., Jaradat S.A., Pantano S., Nagakura R., Boheler K.R.,
 RA Taub D., Hodes R.J., Longo D.L., Schlössinger D., Keller J., Klotz E.,
 RA Kelsoe G., Umezawa A., Vescovi A.L., Rossant J., Kunath T.,
 RA Hogan B.L., Curci A., D'Urso M., Kelso J., Hide W., Ko M.S.;
 RT "Transcriptome analysis of mouse stem cells and early embryos."
 RL PLOS Biol. 1:410-419(2003).
 DR EMBL; AY512913; AAR87784.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 133 AA; 14583 MW; 5929328D1A4B4BF7 CRC64;
 Query Match 46.9%; Score 52.5; DB 2; Length 133;
 Best Local Similarity 60.0%; Pred. No. 20;
 Matches 9; Conservative 1; Mismatches 2; Indels 3; Gaps 1;
 Qy 2 VCRCLCRGVCRCLC 16
 Db 57 VCMCLC---VCLVCV 58
 RESULT 35
 AAR87784 PRELIMINARY; PRT; 133 AA.
 AC AAR87784;
 DT 20-MAY-2004 (TrEMBLrel. 27, Created)
 DT 20-MAY-2004 (TrEMBLrel. 27, Last sequence update)
 DT 20-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Sharov A.A., Piao Y., Matoba R., Dudekula D.B., Qian Y., VanBuren V.,
 RA Falco G., Martin P.R., Stagg C.A., Bassey U.C., Wang Y., Carter M.G.,
 RA Hamatani T., Aiba K., Akutsu H., Sharova L., Tanaka T.S., Kimber W.L.,
 RA Yoshikawa T., Jaradat S.A., Pantano S., Nagakura R., Boheler K.R.,
 RA Taub D., Hodes R.J., Longo D.L., Schlössinger D., Keller J., Klotz E.,
 RA Kelsoe G., Umezawa A., Vescovi A.L., Rossant J., Kunath T.,
 RA Hogan B.L., Curci A., D'Urso M., Kelso J., Hide W., Ko M.S.;
 RT "Transcriptome analysis of mouse stem cells and early embryos."
 RL PLOS Biol. 1:410-419(2003).
 DR EMBL; AY512913; AAR87784.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 133 AA; 14583 MW; 5929328D1A4B4BF7 CRC64;
 Query Match 46.9%; Score 52.5; DB 2; Length 133;
 Best Local Similarity 60.0%; Pred. No. 20;
 Matches 9; Conservative 1; Mismatches 2; Indels 3; Gaps 1;
 Qy 2 VCRCLCRGVCRCLC 16
 Db 57 VCMCLC---VCLVCV 58
 RESULT 36
 Q8BPC0 PRELIMINARY; PRT; 146 AA.
 AC Q8BPC0;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
 DE RIKEN full-length enriched library, clone:5031438A03

DE product:hypothetical protein, full insert sequence.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RA The FANTOM Consortium;
 RT the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Akiyama J., Nishi K., Kitzunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer."
 RL Genome Res. 10:1757-1771(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitzunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer."
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka I., Kasukawa T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakanura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK077267; BAC36719.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 146 AA; 16179 MW; 711E0C93BC80DBBB CRC64;
 Query Match 46.9%; Score 52.5; DB 2; Length 146;
 Best Local Similarity 52.9%; Pred. No. 22;
 Matches 9; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

QY 2 VCRCLRRGVCRCLRR 18
 DB 96 VCMVCVC---VCVCVCMR 109

RESULT 37

Q9H654 PRELIMINARY; PRT; 160 AA.

AC Q9H654
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein FLJ22596.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

SEQUENCE FROM N.A.

RC TISSUE=Human small intestine;
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Ohayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isoigai T., Sugano S.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK026249; BAB15412.1; -.
 SQ SEQUENCE 160 AA; 16738 MW; FBB334293C8E27E7 CRC64;

Query Match 46.9%; Score 52.5; DB 2; Length 160;
 Best Local Similarity 56.2%; Pred. No. 24;
 Matches 9; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 1 GVCRCRCRRGVCRCLC 16
 DB 34 GVCVCVC---VCVCVC 46

RESULT 38

Q6GQP2 PRELIMINARY; PRT; 462 AA.

AC Q6GQP2
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]

SEQUENCE FROM N.A.

RC TISSUE=Whole;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RA Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC072700; AAH2700.1; -.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 462 AA; 52504 MW; 9BF8B5754E98355A CRC64;

Query Match 46.9%; Score 52.5; DB 2; Length 462;
 Best Local Similarity 56.2%; Pred. No. 60;
 Matches 9; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 1 GVCRCRCRRGVCRCLC 16
 DB 419 GVCVCVC---VCVCVC 431

RESULT 39

Q9XVX3 PRELIMINARY; PRT; 152 AA.

AC Q9XVX3
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein C06A1.6.
 GN Name=C06A1.6;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]

SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology.";
 RL Science 282:2012-2018 (1998).
 RN [2]

SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;
 RA McMurray A.A.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z49886; CAA90055.1; -.
 DR PIR; T18975; T18975.
 DR HSP; P02876; 9WGA.

DR WormPep; C06A1.6; CE20484.

DR InterPro; IPR001450; 4Fe4s ferredoxin.

DR InterPro; IPR006209; EGF like.

DR InterPro; IPR004035; EndoIII_FCL.

DR InterPro; IPR001007; WFC.

DR PROSITE; PS00199; 4Fe4s FERREDOXIN; UNKNOWN_1.

DR PROSITE; PS00022; EGF_1; UNKNOWN_1.

DR PROSITE; PS00764; ENDONUCLEASE_III_1; UNKNOWN_1.

DR PROSITE; PS01208; WFC_1; 1.

KW Hypothetical protein.

SQ SEQUENCE 152 AA; 15646 MW; 6E254F0BEA76D354 CRC64;

Query Match 46.4%; Score 52; DB 2; Length 152;
 Best Local Similarity 52.9%; Pred. No. 26;
 Matches 9; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GVCRCRCRRGVCRCLC 17
 DB 67 GCGGCCCCPRCCCCCR 83

RESULT 40

Q8IRLO PRELIMINARY; PRT; 1506 AA.

AC Q8IRLO;

01-MAR-2003 (TrEMBLrel. 23, Created)
 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 CG32681-PA.
 ORFNames=CG32681;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 SEQUENCE FROM N.A.
 RP MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Beres P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Flesher C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 [2]
 SEQUENCE FROM N.A.
 RP MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
 melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
 [3]
 SEQUENCE FROM N.A.
 RP MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskaas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
 [4]

RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochownik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
 [5]
 SEQUENCE FROM N.A.
 RP FLYBASE;
 RG Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 SEQUENCE FROM N.A.
 RP FLYBASE;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003450; AA092801; -;
 DR FLYBASE; FB00052681; CG32681.
 SQ SEQUENCE 1506 AA; 169457 MW; 13193417CFAA9772 CRC64;
 Query Match 46.4%; Score 52; DB 2; Length 1506;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 4 RCLCRRGVCRCLCRR 18
 DB 199 RCPSSRGSCSLKRR 213
 Search completed: October 26, 2004, 15:39:22
 Job time : 130.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 15:13:05 ; Search time 115 Seconds

(without alignments)
56.149 Million cell updates/sec

Title: US-10-009-317A-33

Perfect score: 112

Sequence: 1 GVCRLCRRGVCRLCRR 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Seq04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 112 | 100.0 | 18 | 4 | AAB35047 Theta def |
| 2 | 112 | 100.0 | 18 | 5 | ABP53296 Anti-vira |
| 3 | 112 | 100.0 | 18 | 8 | ADO35231 Rhesus th |
| 4 | 112 | 100.0 | 18 | 8 | ADO35242 Rhesus th |
| 5 | 112 | 100.0 | 18 | 8 | ADO35241 Rhesus th |
| 6 | 107 | 95.5 | 18 | 8 | ADO35243 Rhesus th |
| 7 | 102 | 91.1 | 18 | 5 | ABP53298 Anti-vira |
| 8 | 99 | 88.4 | 18 | 4 | AAB35030 Theta def |
| 9 | 99 | 88.4 | 18 | 5 | ABP53297 Anti-vira |
| 10 | 99 | 88.4 | 18 | 6 | AAB33866 Macaca mu |
| 11 | 99 | 88.4 | 18 | 7 | ADP95202 Cyclic de |
| 12 | 99 | 88.4 | 18 | 8 | ADP5357 Antimicro |
| 13 | 99 | 88.4 | 18 | 8 | ADG70012 Rhesus th |
| 14 | 99 | 88.4 | 18 | 8 | ADO35229 Rhesus th |
| 15 | 99 | 88.4 | 18 | 8 | ADO35238 Rhesus th |
| 16 | 99 | 88.4 | 18 | 8 | ADO35239 Rhesus th |
| 17 | 99 | 88.4 | 18 | 8 | ADO35256 Rhesus th |
| 18 | 99 | 88.4 | 38 | 8 | ADO35253 Monkey RT |
| 19 | 95 | 84.8 | 18 | 4 | AAB35037 Rhesus ma |
| 20 | 92 | 82.1 | 18 | 5 | ABP53294 Synthetic |
| 21 | 92 | 82.1 | 18 | 6 | AAB33801 Human ret |
| 22 | 92 | 82.1 | 18 | 6 | AAB33863 Enantio-r |
| 23 | 92 | 82.1 | 18 | 8 | ADN08176 Human ret |
| 24 | 89 | 79.5 | 18 | 5 | ABP53299 Anti-vira |
| 25 | 89 | 79.5 | 18 | 6 | AAB33804 I15Y retr |

| | | | | | | |
|----|----|------|----|---|----------|--------------------|
| 26 | 89 | 79.5 | 18 | 6 | AAB33864 | Aae33864 Enantio-R |
| 27 | 89 | 79.5 | 18 | 6 | AAB33803 | Aae33803 I6Y retro |
| 28 | 89 | 79.5 | 18 | 6 | AAB33802 | Aae33802 R9K retro |
| 29 | 89 | 79.5 | 18 | 8 | ADO35244 | Ado35244 Rhesus th |
| 30 | 89 | 79.5 | 18 | 8 | ADN08177 | Adn08177 Human ret |
| 31 | 89 | 79.5 | 18 | 8 | ADN08179 | Adn08179 Human ret |
| 32 | 89 | 79.5 | 18 | 8 | ADN08178 | Adn08178 Human ret |
| 33 | 88 | 78.6 | 18 | 6 | AAB33805 | Aae33805 I2Y retro |
| 34 | 88 | 78.6 | 18 | 6 | AAB33806 | Aae33806 I1Y retr |
| 35 | 88 | 78.6 | 18 | 8 | ADN08180 | Adn08180 Human ret |
| 36 | 88 | 78.6 | 18 | 8 | ADN08181 | Adn08181 Human ret |
| 37 | 86 | 76.8 | 18 | 4 | AAB35046 | Abp35046 Theta def |
| 38 | 86 | 76.8 | 18 | 5 | ABP53295 | Abp53295 Anti-vira |
| 39 | 86 | 76.8 | 18 | 6 | AAB33865 | Aae33865 RC-101/10 |
| 40 | 86 | 76.8 | 18 | 8 | ADO35230 | Ado35230 Rhesus th |
| 41 | 86 | 76.8 | 18 | 8 | ADO35255 | Ado35255 Rhesus th |
| 42 | 86 | 76.8 | 18 | 8 | ADO35249 | Ado35249 Rhesus th |
| 43 | 85 | 75.9 | 18 | 6 | AAB33807 | Aae33807 R4Y retro |
| 44 | 85 | 75.9 | 18 | 8 | ADN08182 | Adn08182 Human ret |
| 45 | 84 | 75.0 | 18 | 8 | ADO35254 | Ado35254 Rhesus th |

ALIGNMENTS

RESULT 1

AAB35047

ID AAB35047 standard; peptide; 18 AA.

XX

AC AAB35047;

XX

XX 27-MAR-2001 (first entry)

DT

XX

XX Theta defensin SEQ ID NO: 31.

DE

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

RESULT 2
ID ABP53296 standard; peptide; 18 AA.
AC ABP53296;
XX 13-NOV-2002 (first entry)
DT
DE Anti-viral theta defensin peptide RTD-3 SEQ ID NO:29.
XX
XX Anti-viral; viral infection; theta-defensin; lipid environment;
KW amphipathic alpha-helical structure; virucide; anti-HIV; immunisation;
KW viral growth inhibitor; viral proliferation inhibitor.
XX
XX Macaca mulatta.
OS Synthetic.
XX WO200260468-A2.
FN
XX 08-AUG-2002.
XX
XX 29-JAN-2002; 2002WO-US002435.
XX
XX 30-JAN-2001; 2001US-0265270P.
PR
XX 01-AUG-2001; 2001US-0309368P.
XX
XX (IOWA) UNIV IOWA RES FOUND.
XX
XX Maury W, Scapleton J, Stinski M, Roller R, Mccray PB, Tack B;
PI WPI; 2002-674815/72.
XX
XX
XX New method of using a first anti-viral peptide comprising a Theta-
PT defensin peptide in an amphipathic Alpha-helical structure in a lipid
PT environment for reducing the infectivity of a virus.
XX
XX Disclosure; Page 10; 65pp; English.
XX
XX The present invention describes a method (M1) of using a first anti-viral
CC peptide (I) comprising a theta-defensin peptide in an amphipathic alpha-
CC helical structure in a lipid environment for reducing the infectivity of
CC a virus. (I) can have virucide and anti-HIV activities, and can be used
CC to reduce virus growth, infectivity burden, shed, and development of anti
CC -viral resistance. (I) can be used for inhibiting the growth and
CC proliferation of a virus and so can be used for; (a) protecting or
CC treating subject from a viral infection, preventing recurrent viral
CC infection in a subject harbouring a latent virus, controlling virus
CC spread within a virally-infected subject (VS), reducing viral burden in a
CC VS, reducing virus shed from a VS, reducing percentage of VS in a
CC population regardless of viral infection status, or inducing latency in a
CC VS; (b) reducing the infectivity of a virus; and (c) rendering virus-
CC contaminated tissue or fluid sample safe for use, or reducing the number
CC of infectious virus particles in a population of viruses. (M1) is useful
CC for reducing the infectivity of a virus in sheep, cattle, horses, swine,
CC cats, fowl and humans e.g. an enveloped virus infecting humans such as
CC human immunodeficiency virus (HIV). Preferably, the anti-viral peptide is
CC administered to a patient who is immunosuppressed or to a subject who is
CC not infected with the virus, where the first anti-viral peptide is
CC administered prior to or subsequent to the virus contacting the subject.
CC The anti-viral peptide is most preferably administered to a subject who
CC is chronically, latently or acutely infected with the virus. The present
CC sequence represents a rhesus monkey theta defensin anti-viral peptide,
CC which is given in the exemplification of the present invention
XX
XX Sequence 18 AA;

Query Match 100.0%; Score 112; DB 5; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.2e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GVCRCLCRGVCRCLCR 18

1 GVCRCLCRGVCRCLCR 18

Db
RESULT 3
ID ADO35231 standard; peptide; 18 AA.
XX
XX ADO35231;
AC
XX 15-JUL-2004 (first entry)
DT
XX Rhesus theta defensin peptide, RTD-3.
DE
XX Monkey; Rhesus theta defensin; RTD-3; antimicrobial peptide; cyclic;
KW antimicrobial; antiinflammatory; antibacterial; virucide; fungicide;
KW food; contact lens solution; eye wash solution; inflammatory response;
KW microbicidal inhibition; microbistatic growth inhibition; disinfectant;
KW food preservative; bacterial infection; viral infection;
KW fungal infection; haemolytic activity.
XX
XX Macaca mulatta.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 1..18
FT Note= "The peptide is cyclised by a covalent link
FT between these two residues"
FT Disulfide-bond 3..16
FT Disulfide-bond 5..14
FT Disulfide-bond 7..12
XX
XX US2004014669-A1.
FN
XX 22-JAN-2004.
PD
XX 30-APR-2003; 2003US-00427715.
PF
XX 30-APR-2002; 2002US-0377071P.
PR
XX (REGC) UNIV CALIFORNIA.
PA
XX Selsted ME, Tran DQ;
PI WPI; 2004-167945/16.
XX
XX Novel theta defensin analog useful for reducing or inhibiting growth or
PT survival of a microorganism in an environment such as food or food
PT product, contact lens solution, or eye wash solution, an inanimate
PT object.
XX
XX Example 1; SEQ ID NO 3; 46pp; English.
XX
XX The invention relates to a theta defensin analogue defined by formulae
CC detailed in the claims or appearing as ADO35231-ADO35257. The theta
CC defensin analogue is useful for reducing or inhibiting growth or survival
CC of a microorganism in an environment capable of sustaining the growth or
CC survival of the microorganism and is useful for reducing or inhibiting
CC growth or survival of a microorganism in an environment such as food or
CC food product, a solution (e.g., contact lens solution, or eye wash
CC solution), an inanimate object comprising surface, or a mammal. The
CC peptides are also useful for decreasing inflammatory response and for
CC microbicidal inhibition of survival of microorganism as well as
CC microbistatic inhibition of growth. Thus the peptides are useful as
CC therapeutic agents, disinfectants, food preservatives, or medicaments.
CC The peptides are also useful for treating a patient suffering from
CC bacterial, viral, fungal or other infection. The theta defensins have
CC high antimicrobial activity and low haemolytic activity. The present
CC sequence represents the rhesus monkey wild-type theta defensin RTD-3.
XX
XX Sequence 18 AA;
Query Match 100.0%; Score 112; DB 8; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.2e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | |
|----------|---|-------------------------------------|
| | Query Match | 100.0%; Score 112; DB 8; Length 18; |
| | Best Local Similarity | 100.0%; Pred. NO. 9.2e-06; |
| | Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| QY | 1 GVCRCLCRGGVCKLCRR 18 | |
| | | |
| DD | 1 GVCRCLCRGGVCKLCRR 18 | |
| | | |
| RESULT 5 | | |
| AD035241 | | |
| ID | AD035241 standard; peptide; 18 AA. | |
| XX | AD035241; | |
| AC | | |
| XX | | |
| DT | 15-JUL-2004 (first entry) | |
| DE | Rhesus theta defensin analogue peptide arTPD-3-OH. | |
| XX | | |
| KW | Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial; | |
| KW | anti-inflammatory; antibacterial; virucide; fungicide; food; | |
| KW | contact lens solution; eye wash solution; inflammatory response; | |
| KW | microbicidal inhibition; microbistatic growth inhibition; disinfectant; | |
| KW | food preservative; bacterial infection; viral infection; | |
| KW | fungal infection; haemolytic activity. | |
| XX | | |
| OS | Macaca mulatta. | |
| OS | Synthetic. | |
| XX | | |
| FFH | Key Location/Qualifiers | |
| FT | Disulfide-bond 3..16 | |
| FT | Disulfide-bond 5..14 | |
| FT | Disulfide-bond 7..12 | |
| FT | Modified-site 18 | |
| FT | /note= "Hydroxylated" | |
| XX | | |
| PN | US2004014669-A1. | |
| XX | | |
| PD | 22-JAN-2004. | |
| XX | | |
| PF | 30-APR-2003; 2003US-00427715. | |
| XX | | |
| PR | 30-APR-2002; 2002US-0377071P. | |
| XX | (REGC) UNIV CALIFORNIA. | |
| PA | | |
| PI | Selsted ME, Tran DQ; | |
| XX | | |
| DR | WPI; 2004-167945/16. | |
| XX | | |
| PPT | Novel theta defensin analog useful for reducing or inhibiting growth or | |
| PPT | survival of a microorganism in an environment such as food or food | |
| PPT | product, contact lens solution, or eye wash solution, an inanimate | |
| PPT | object. | |
| PT | | |
| XX | | |
| PS | Claim 1; SEQ ID NO 15; 46pp; English. | |
| XX | | |
| CC | The invention relates to a theta defensin analogue defined by formulae | |
| CC | detailed in the claims or appearing as AD035239-ADO35257. The theta | |
| CC | defensin analogue is useful for reducing or inhibiting growth or survival | |
| CC | of a microorganism in an environment capable of sustaining the growth or | |
| CC | survival of the microorganism and is useful for reducing or inhibiting | |
| CC | growth or survival of a microorganism in an environment such as food or | |
| CC | food product, a solution (e.g., contact lens solution, or eye wash | |
| CC | solution), an inanimate object comprising surface, or a mammal. The | |
| CC | peptides are also useful for decreasing inflammatory response and for | |
| CC | microbicidal inhibition of survival of microorganism as well as | |
| CC | microbistatic inhibition of growth. Thus the peptides are useful as | |
| CC | therapeutic agents, disinfectants, food preservatives, or medicaments. | |
| CC | The peptides are also useful for treating a patient suffering from | |
| CC | bacterial, viral, fungal or other infection. The theta defensins have | |
| CC | high antimicrobial activity and low haemolytic activity. The present | |

CC sequence represents a Rhesus theta defensin analogue peptide.

XX Sequence 18 AA;
SQ Query Match 100.0%; Score 112; DB 8; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.2e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GVCRLCRRGVGRCCLRR 18
Db 1 GVCRLCRRGVGRCCLRR 18

RESULT 6
ADO35243
ID ADO35243 standard; peptide; 18 AA.

XX AC ADO35243;

XX DT 15-JUL-2004 (first entry)

XX DE Rhesus theta defensin analogue peptide 3:1 arTD-1-NH.

XX KW Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial;
XX KW antiinflammatory; antibacterial; virucide; fungicide; food;
XX KW contact lens solution; eye wash solution; inflammatory response;
XX KW microbicidal inhibition; microbistatic growth inhibition; disinfectant;
XX KW food preservative; bacterial infection; viral infection;
XX KW fungal infection; haemolytic activity.

XX OS Macaca mulatta.

XX OS Synthetic.

XX PH Key Location/Qualifiers

XX FT Disulfide-bond 4. 17

XX FT Disulfide-bond 6. 15

XX FT Disulfide-bond 8. 13

XX FT Modified-site 18

XX FT /note= "Amidated"

XX US2004014669-A1.

XX PD 22-JAN-2004.

XX PF 30-APR-2003; 2003US-00427715.

XX PR 30-APR-2002; 2002US-0377071P.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Selsted ME, Tran DQ;

XX DR WPI; 2004-167945/16.

XX PT Novel theta defensin analog useful for reducing or inhibiting growth or
PT survival of a microorganism in an environment such as food or food
PT product, contact lens solution, or eye wash solution, an inanimate
PT object.

XX PS Claim 1; SEQ ID NO 17; 46pp; English.

XX CC The invention relates to a theta defensin analogue defined by formulae
CC detailed in the claims or appearing as ADO35239-ADO35257. The theta
CC defensin analogue is useful for reducing or inhibiting growth or survival
CC of a microorganism in an environment capable of sustaining the growth or
CC survival of the microorganism and is useful for reducing or inhibiting
CC growth or survival of a microorganism in an environment such as food or
CC food product, a solution (e.g., contact lens solution, or eye wash
CC solution), an inanimate object comprising surface, or a mammal. The
CC peptides are also useful for decreasing inflammatory response and for
CC microbicidal inhibition of survival of microorganism as well as
CC microbistatic inhibition of growth. Thus the peptides are useful as
CC therapeutic agents, disinfectants, food preservatives, or medicaments.

CC The peptides are also useful for treating a patient suffering from
CC bacterial, viral, fungal or other infection. The theta defensins have
CC high antimicrobial activity and low haemolytic activity. The present
CC sequence represents a Rhesus theta defensin analogue peptide.

XX SQ Sequence 18 AA;

Query Match 95.5%; Score 107; DB 8; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVCRLCRRGVGRCCLCR 17

Db 2 GVCRLCRRGVGRCCLCR 18

RESULT 7

ABP53298

ID ABP53298 standard; peptide; 18 AA.

XX AC ABP53298;

XX DT 13-NOV-2002 (first entry)

XX DE Anti-viral chimeric theta defensin peptide H/RTD-3 SEQ ID NO:31.

XX KW Anti-viral; viral infection; theta-defensin; lipid environment;

XX KW amphipathic alpha-helical structure; virucide; anti-HIV; immunisation;

XX KW viral growth inhibitor; viral proliferation inhibitor.

XX OS Homo sapiens.

XX OS Macaca mulatta.

XX OS Synthetic.

XX PN WO200260468-A2.

XX PD 08-AUG-2002.

XX PF 29-JAN-2002; 2002WO-US002435.

XX PR 30-JAN-2001; 2001US-0265270P.

XX PR 01-AUG-2001; 2001US-0309368P.

XX PA (IOWA) UNIV IOWA RES FOUND.

XX PI Maury W, Stapleton J, Stinski M, Roller R, Mccray PB, Tack B;

XX DR WPI; 2002-674815/72.

XX PT New method of using a first anti-viral peptide comprising a Theta-
PT defensin peptide in an amphipathic Alpha-helical structure in a lipid
PT environment for reducing the infectivity of a virus.

XX PS Disclosure; Page 10; 65pp; English.

XX CC The present invention describes a method (M1) of using a first anti-viral
CC peptide (I) comprising a theta-defensin peptide in an amphipathic alpha-
CC helical structure in a lipid environment for reducing the infectivity of
CC a virus. (I) can have virucide and anti-HIV activities, and can be used
CC to reduce virus growth, infectivity burden, shed, and development of anti-
CC viral resistance. (I) can be used for inhibiting the growth and
CC proliferation of a virus and so can be used for; (a) protecting or
CC treating subject from a viral infection, preventing recurrent viral
CC infection in a subject harbouring a latent virus, controlling virus
CC spread within a virally-infected subject (VS), reducing viral burden in a
CC VS, reducing virus shed from a VS, reducing percentage of VS in a
CC population regardless of viral infection status, or inducing latency in a
CC VS; (b) reducing the infectivity of a virus; and (c) rendering the number-
CC contaminated tissue or fluid sample safe for use, or reducing the number
CC of infectious virus particles in a population of viruses. (M1) is useful
CC for reducing the infectivity of a virus in sheep, cattle, horses, swine,
CC cats, fowl and humans e.g. an enveloped virus infecting humans such as
CC human immunodeficiency virus (HIV). Preferably, the anti-viral peptide is

Qy 1 GVCRLCRRGVCRLCRR 18

Query Match 88.4%; Score 99; DB 5; Length 18;
Best Local Similarity 83.3%; Pred. No. 0.00023;
Matches 15: Conservative 1; Mismatches 2; Indels

OV 1 GVCRCLCRRGVCRCLCRR 18
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Best Local Similarity 83.3%; Pred. No. 0.00023;

ADD35357
 ID ADD35357 standard; peptide; 18 AA.
 XX
 AC ADD35357;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Antimicrobial peptide theta-defensin.
 XX
 DE antimicrobial; opthalmic; prostaglandin; hypotensive; ophthalmological;
 KW intraocular pressure; glaucoma; ocular hypertension; hyperaemia;
 KW irritation; inflammation; conjunctiva; ocular cell dysplasia;
 KW iridial melanocyte hyperplasia; hyperpigmentation.
 XX
 OS Unidentified.
 XX
 XX WO2003079997-A2.
 PN
 XX
 PD 02-OCT-2003.
 XX
 XX
 PF 21-MAR-2003; 2003WO-US008935.
 XX
 XX 21-MAR-2002; 2002US-0367071P.
 PR
 XX (CAYM-) CAYMAN CHEM CO.
 PA
 XX
 PI Maxey KM, Johnson J;
 XX
 XX WPI; 2004-011506/01.
 DR
 XX
 XX Ophthalmic solution useful for the treatment of increased intraocular
 PT pressure comprises a prostaglandin of the F-series and an antimicrobial
 PT peptide.
 XX
 PS Disclosure; Page 11; 11pp; English.
 XX
 XX The invention relates to a novel ophthalmic solution comprising a
 CC prostaglandin of the F-series and an antimicrobial peptide. A solution of
 CC the invention has hypotensive and ophthalmological activity. The solution
 CC is useful for the treatment of increased intraocular pressure, such as
 CC caused by glaucoma and for the reduction of ocular hypertension. The
 CC prostaglandin and the antimicrobial peptide work synergistically, to
 CC provide beneficial reduction in the incidence of irritant and toxic side
 CC effects such as hyperaemia, irritation and inflammation of conjunctiva,
 CC ocular cell dysplasia, iridial melanocyte hyperplasia, and
 CC hyperpigmentation, associated with the prior art prostaglandin
 CC compositions. The present sequence represents an antimicrobial peptide of
 CC the invention.
 XX
 XX Sequence 18 AA;
 SQ
 Query Match 88.4%; Score 99; DB 8; Length 18;
 Best Local Similarity 83.3%; Pred. No. 0.00023;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GVCRCCLCRGVCRCLCRR 18
 Db 1 GFCRCCLCRGVCRCLCCTR 18
 RESULT 13
 ADG70012
 ID ADG70012 standard; peptide; 18 AA.
 XX
 XX
 AC ADG70012;
 XX
 XX 11-MAR-2004 (first entry)
 DT
 XX
 XX Rhesus theta-defensin-1 (RTD-1) peptide.
 DE
 XX
 XX rhesus theta defensin-1; RTD-1; bacteraemia; lipopolysaccharide; LPS;
 KW lipoteichoic acid; LTA; septic shock; antibacterial; fungicide; virucide;
 KW immunomodulator; anticoagulant activity;

KW microbial cell-wall biosynthesis; immunomodulation; anticoagulant.
 XX
 OS Macaca mulatta.
 XX
 XX WO2003105883-A1.
 PN
 XX 24-DEC-2003.
 PD
 XX 30-MAY-2003; 2003WO-EP005694.
 PF
 XX 13-JUN-2002; 2002DE-01026216.
 PR
 XX (FARB) BAYER HEALTHCARE AG.
 PA
 XX
 XX Ladel C, Newton B, Labischinski H, Brunner N, Gerdes C;
 PI
 XX WPI; 2004-071500/07.
 DR
 XX
 XX Use of rhesus theta defensin-1 for treating or preventing bacteraemia and
 PT septic shock, also for binding bacterial products and as immunomodulator
 PT and anticoagulant.
 XX
 XX Example 1; SEQ ID NO 1; 28pp; German.
 PS
 XX This invention describes the novel use of rhesus theta defensin-1 (RTD-1)
 CC for preparing a composition for treatment and/or prevention of
 CC bacteraemia for binding bacterial products such as lipopolysaccharide
 CC (LPS) and/or lipoteichoic acid (LTA), or for treatment and/or prevention
 CC of septic shock. RTD-1, isolated from immune cells of rhesus monkeys, has
 CC antibacterial, fungicide, virucide, immunomodulator and anticoagulant
 CC activity. RTD-1 inhibits microbial cell-wall biosynthesis and also binds
 CC to LPS and LTA. RTD-1 is useful for treatment and prevention of severe
 CC infections caused by Gram-positive or -negative bacteria and yeasts, or
 CC by viruses. RTD-1 combines four advantageous properties: a direct
 CC antimicrobial action, neutralisation of bacterial products (by binding),
 CC immunomodulation (reducing release of proinflammatory cytokines but
 CC increasing release of regulatory factors) and anticoagulant action, so
 CC provides a better and simpler treatment.
 XX
 XX Sequence 18 AA;
 SQ
 Query Match 88.4%; Score 99; DB 8; Length 18;
 Best Local Similarity 83.3%; Pred. No. 0.00023;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GVCRCCLCRGVCRCLCRR 18
 Db 1 GFCRCCLCRGVCRCLCCTR 18
 RESULT 14
 ADO35229
 ID ADO35229 standard; peptide; 18 AA.
 XX
 XX ADO35229;
 AC
 XX 15-JUL-2004 (first entry)
 DT
 XX
 XX Rhesus theta defensin peptide, RTD-1.
 DE
 XX
 XX Monkey; Rhesus theta defensin; RTD-1; antimicrobial peptide; cyclic;
 KW antimicrobial; antiinflammatory; antibacterial; virucide; fungicide;
 KW food; contact lens solution; eye wash solution; inflammatory response;
 KW microbicidal inhibition; microbistatic growth inhibition; disinfectant;
 KW food preservative; bacterial infection; viral infection;
 KW fungal infection; haemolytic activity.
 XX
 XX Macaca mulatta.
 OS
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1..18
 FT /note= "The peptide is cyclised by a covalent link
 FT between these two residues"

FT Disulfide-bond 3. .16
 FT Disulfide-bond 5. .14
 FT Disulfide-bond 7. .12
 XX
 PN US2004014669-A1.
 XX
 PD 22-JAN-2004.
 XX

XX 30-APR-2003; 2003US-00427715.
 XX
 XX 30-APR-2002; 2002US-0377071P.
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX Selsted ME, Tran DQ;
 PI
 XX WPI; 2004-167945/16.
 DR

XX Novel theta defensin analog useful for reducing or inhibiting growth or
 XX survival of a microorganism in an environment such as food or food
 XX product, contact lens solution, or eye wash solution, an inanimate
 XX object.
 PT
 XX Example 1; SEQ ID NO 1; 46pp; English.

XX The invention relates to a theta defensin analogue defined by formulae
 XX detailed in the claims or appearing as ADO35239-ADO35257. The theta
 XX defensin analogue is useful for reducing or inhibiting growth or survival
 XX of a microorganism in an environment capable of sustaining the growth or
 XX survival of the microorganism and is useful for reducing or inhibiting
 XX growth or survival of a microorganism in an environment such as food or
 XX food product, a solution (e.g., contact lens solution, or eye wash
 XX solution), an inanimate object comprising surface, or a mammal. The
 XX peptides are also useful for decreasing inflammatory response and for
 XX microbicidal inhibition of survival of microorganism as well as
 XX microbistatic inhibition of growth. Thus the peptides are useful as
 XX therapeutic agents, disinfectants, food preservatives, or medicaments.
 XX The peptides are also useful for treating a patient suffering from
 XX bacterial, viral, fungal or other infection. The theta defensins have
 XX high antimicrobial activity and low haemolytic activity. The present
 XX sequence represents the rhesus monkey wild-type theta defensin RDT-1.

XX Sequence 18 AA;
 SQ

Query Match 88.4%; Score 99; DB 8; Length 18;
 Best Local Similarity 83.3%; Pred. No. 0.00023;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 GVCRCCLCRGVCRCICRR 18
 Db 1 GFCRCCLCRGVCRCICTR 18
 | | | | | | | | | | | | | | | | | |
 1 1

RESULT 15
 ADO35238

ID ADO35238 standard; peptide; 18 AA.
 XX
 AC ADO35238;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 XX Rhesus theta defensin analogue peptide aRTD-1-OH.

XX Monkey; Rhesus theta defensin; RDT; antimicrobial peptide; antimicrobial;
 XX antiinflammatory; antibacterial; virucide; fungicide; food;
 XX contact lens solution; eye wash solution; inflammatory response;
 XX microbicidal inhibition; microbistatic growth inhibition; disinfectant;
 XX food preservative; bacterial infection; viral infection;
 XX fungal infection; haemolytic activity.
 OS Macaca mulatta.
 OS Synthetic.
 XX

Key Location/Qualifiers
 Disulfide-bond 3. .16
 Disulfide-bond 5. .14
 Disulfide-bond 7. .12
 Modified-site 18
 /note= "Hydroxylated"

US2004014669-A1.

22-JAN-2004.

30-APR-2003; 2003US-00427715.

30-APR-2002; 2002US-0377071P.

(REGC) UNIV CALIFORNIA.

Selsted ME, Tran DQ;

WPI; 2004-167945/16.

Novel theta defensin analog useful for reducing or inhibiting growth or
 survival of a microorganism in an environment such as food or food
 product, contact lens solution, or eye wash solution, an inanimate
 object.

Example 2; SEQ ID NO 12; 46pp; English.

The invention relates to a theta defensin analogue defined by formulae
 detailed in the claims or appearing as ADO35239-ADO35257. The theta
 defensin analogue is useful for reducing or inhibiting growth or survival
 of a microorganism in an environment capable of sustaining the growth or
 survival of the microorganism and is useful for reducing or inhibiting
 growth or survival of a microorganism in an environment such as food or
 food product, a solution (e.g., contact lens solution, or eye wash
 solution), an inanimate object comprising surface, or a mammal. The
 peptides are also useful for decreasing inflammatory response and for
 microbicidal inhibition of survival of microorganism as well as
 microbistatic inhibition of growth. Thus the peptides are useful as
 therapeutic agents, disinfectants, food preservatives, or medicaments.
 The peptides are also useful for treating a patient suffering from
 bacterial, viral, fungal or other infection. The theta defensins have
 high antimicrobial activity and low haemolytic activity. The present
 sequence represents a Rhesus theta defensin analogue peptide.

Sequence 18 AA;
 SQ

Query Match 88.4%; Score 99; DB 8; Length 18;
 Best Local Similarity 83.3%; Pred. No. 0.00023;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GVCRCCLCRGVCRCICRR 18
 Db 1 GFCRCCLCRGVCRCICTR 18
 | | | | | | | | | | | | | | | | | |
 1 1

RESULT 16

ADO35239

ID ADO35239 standard; peptide; 18 AA.

AC ADO35239;

DT 15-JUL-2004 (first entry)

Rhesus theta defensin analogue peptide aRTD-1-NH.

Monkey; Rhesus theta defensin; RDT; antimicrobial peptide; antimicrobial;
 antiinflammatory; antibacterial; virucide; fungicide; food;
 contact lens solution; eye wash solution; inflammatory response;
 microbicidal inhibition; microbistatic growth inhibition; disinfectant;
 food preservative; bacterial infection; viral infection;
 fungal infection; haemolytic activity.

XX

```

OS Macaca mulatta.
OS Synthetic.
FH Key Location/Qualifiers
FT Disulfide-bond 3..16
FT Disulfide-bond 5..14
FT Disulfide-bond 7..12
FT Modified-site 18
FT /note= "Amidated"
XX
XX US2004014669-A1.
XX
XX 22-JAN-2004.
XX
XX 30-APR-2003; 2003US-00427715.
XX
XX 30-APR-2002; 2002US-0377071P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Selsted ME, Tran DQ;
XX
XX WPI; 2004-167945/16.
XX
XX Novel theta defensin analog useful for reducing or inhibiting growth or
XX survival of a microorganism in an environment such as food or food
XX product, contact lens solution, or eye wash solution, an inanimate
XX object.
XX
XX Claim 1; SEQ ID NO 13; 46pp; English.
XX
XX The invention relates to a theta defensin analogue defined by formulae
XX detailed in the claims or appearing as ADO35239-ADO35257. The theta
XX defensin analogue is useful for reducing or inhibiting growth or survival
XX of a microorganism in an environment capable of sustaining the growth or
XX survival of the microorganism and is useful for reducing or inhibiting
XX growth or survival of a microorganism in an environment such as food or
XX food product, a solution (e.g., contact lens solution, or eye wash
XX solution), an inanimate object comprising surface, or a mammal. The
XX peptides are also useful for decreasing inflammatory response and for
XX microbicidal inhibition of growth. Thus the peptides are useful as
XX therapeutic agents, disinfectants, food preservatives, or medicaments.
XX The peptides are also useful for treating a patient suffering from
XX bacterial, viral, fungal or other infection. The theta defensins have
XX high antimicrobial activity and low haemolytic activity. The present
XX sequence represents a Rhesus theta defensin analogue peptide.
XX
XX Sequence 18 AA;
XX
XX Query Match 88.4%; Score 99; DB 8; Length 18;
XX Best Local Similarity 83.3%; Pred. No. 0.00023;
XX Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 GVCRCRCRGVCRCLCRR 18
XX | ||||| ||||| ||||| |||||
XX Db 1 GFCRCRCRGVCRCLCTR 19
XX
XX RESULT 17
XX ID ADO35256
XX XX ADO35256 standard; peptide; 18 AA.
XX
XX AC ADO35256;
XX
XX DT 15-JUL-2004 (first entry)
XX
XX Rhesus theta defensin analogue peptide RDT-5.
XX
XX Monkey; Rhesus theta defensin; RDT; antimicrobial peptide; antimicrobial;
XX antiinflammatory; antibacterial; virucide; fungicide; food;
XX contact lens solution; eye wash solution; inflammatory response;
XX microbicidal inhibition; microbistatic growth inhibition; disinfectant;

```

```

KW food preservative; bacterial infection; viral infection;
KW fungal infection; haemolytic activity; cyclic.
XX
XX Macaca mulatta.
OS Synthetic.
FH Key Location/Qualifiers
FT Modified-site 1..18
FT /note= "The peptide is cyclised by a covalent link
FT between these two residues"
FT
FT Disulfide-bond 3..16
FT Disulfide-bond 5..14
FT Disulfide-bond 7..12
XX
XX US2004014669-A1.
XX
XX 22-JAN-2004.
XX
XX 30-APR-2003; 2003US-00427715.
XX
XX 30-APR-2002; 2002US-0377071P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Selsted ME, Tran DQ;
XX
XX WPI; 2004-167945/16.
XX
XX Novel theta defensin analog useful for reducing or inhibiting growth or
XX survival of a microorganism in an environment such as food or food
XX product, contact lens solution, or eye wash solution, an inanimate
XX object.
XX
XX Claim 1; SEQ ID NO 30; 46pp; English.
XX
XX The invention relates to a theta defensin analogue defined by formulae
XX detailed in the claims or appearing as ADO35239-ADO35257. The theta
XX defensin analogue is useful for reducing or inhibiting growth or survival
XX of a microorganism in an environment capable of sustaining the growth or
XX survival of the microorganism and is useful for reducing or inhibiting
XX growth or survival of a microorganism in an environment such as food or
XX food product, a solution (e.g., contact lens solution, or eye wash
XX solution), an inanimate object comprising surface, or a mammal. The
XX peptides are also useful for decreasing inflammatory response and for
XX microbicidal inhibition of survival of microorganism as well as
XX therapeutic agents, disinfectants, food preservatives, or medicaments.
XX The peptides are also useful for treating a patient suffering from
XX bacterial, viral, fungal or other infection. The theta defensins have
XX high antimicrobial activity and low haemolytic activity. The present
XX sequence represents a Rhesus theta defensin analogue peptide.
XX
XX Sequence 18 AA;
XX
XX Query Match 88.4%; Score 99; DB 8; Length 18;
XX Best Local Similarity 87.5%; Pred. No. 0.00023;
XX Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GVCRCRCRGVCRCLC 16
XX | : ||||| ||||| |||||
XX Db 1 GICRCRCRGVCRCLC 16
XX
XX RESULT 18
XX ID ADO35263
XX XX ADO35263 standard; protein; 38 AA.
XX
XX AC ADO35263;
XX
XX DT 15-JUL-2004 (first entry)
XX
XX Monkey RDT-1 (rhesus theta defensin 1) 2X protein.

```

XX Monkey; Rhesus theta defensin; RTD-1; antimicrobial peptide;
 KW antimicrobial; antiinflammatory; antibacterial; virucide; fungicide;
 KW food; contact lens solution; eye wash solution; inflammatory response;
 KW microbicidal inhibition; microbistatic growth inhibition; disinfectant;
 KW food preservative; bacterial infection; viral infection;
 KW fungal infection; haemolytic activity.
 XX
 OS Macaca mulatta.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 XX Cleavage-site 1..2 /note= "CNBR cleavage site"
 XX Peptide 2..19 /note= "RTD-1"
 XX Cleavage-site 20..21 /note= "CNBR cleavage site"
 XX Peptide 21..38 /note= "RTD-1"
 XX
 XX US2004014669-A1.
 XX
 XX 22-JAN-2004.
 XX
 XX 30-APR-2003; 2003US-00427715.
 XX
 XX 30-APR-2002; 2002US-0377071P.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX Selsted ME, Tran DQ;
 XX
 XX WPI: 2004-167945/16.
 XX N-PSDB; ADO35262.
 XX
 XX Novel theta defensin analog useful for reducing or inhibiting growth or
 PT survival of a microorganism in an environment such as food or food
 PT product, contact lens solution, or eye wash solution, an inanimate
 PT object.
 XX
 XX Example 4; Fig 17a; 46pp; English.
 XX
 XX The invention relates to a theta defensin analogue defined by formulae
 CC detailed in the claims or appearing as ADO35239-ADO35257. The theta
 CC defensin analogue is useful for reducing or inhibiting growth or survival
 CC of a microorganism in an environment capable of sustaining the growth or
 CC survival of the microorganism and is useful for reducing or inhibiting
 CC growth or survival of a microorganism in an environment such as food or
 CC food product, a solution (e.g., contact lens solution, or eye wash
 CC solution), an inanimate object comprising surface, or a mammal. The
 CC peptides are also useful for decreasing inflammatory response and for
 CC microbicidal inhibition of survival of microorganism as well as
 CC microbistatic inhibition of growth. Thus the peptides are useful as
 CC therapeutic agents, disinfectants, food preservatives, or medicaments.
 CC The peptides are also useful for treating a patient suffering from
 CC bacterial, viral, fungal or other infection. The theta defensins have
 CC high antimicrobial activity and low haemolytic activity. The present
 CC sequence is protein containing 2 copies of the rhesus monkey theta
 CC defensin, RTD-1 peptide.
 XX
 XX Sequence 38 AA;
 XX
 Query Match 88.4%; Score 99; DB 8; Length 38;
 Best Local Similarity 83.3%; Pred. No. 0.00041;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 1 GVCRCRCRGVCRCLCRR 18
 DB 2 GFCRCRCRGVCRCLCTR 19
 RESULT 19
 AAB35037

ID AAB35037 standard; peptide; 18 AA.
 XX
 AC AAB35037;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Rhesus macaque theta defensin peptide SEQ ID NO: 8.
 XX
 XX Theta defensin; antimicrobial; cyclic; bacterium; fungus; protozoan;
 KW virus; helminth; disinfectant; food preservative; analogue.
 XX
 OS Rhesus macaque.
 XX
 XX WO2000068265-A1.
 XX
 XX 16-NOV-2000.
 XX
 XX 10-MAY-2000; 2000WO-US012842.
 XX
 XX 10-MAY-1999; 99US-00309487..
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX Selsted ME, Tang Y, Yuan J, Quellette AJ;
 XX
 XX WPI; 2001-031853/04.
 XX
 XX Novel theta defensin peptide with antimicrobial activity against
 PT bacteria, yeast, fungi, protozoa and viruses.
 XX
 XX Example 1; Fig 2; 110pp; English.
 XX
 XX The present invention provides theta defensin peptides and analogues
 CC which have antimicrobial activity. They can be used in the treatment of
 CC bacterial, viral, fungal, protozoan and helminthic infections, in
 CC disinfectants and as food preservatives
 XX
 XX Sequence 18 AA;
 XX
 Query Match 84.8%; Score 95; DB 4; Length 18;
 Best Local Similarity 87.5%; Pred. No. 0.00062;
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 GVCRCRCRGVCRCLC 16
 DB 3 GFCRCRCRGVCRCLC 18
 RESULT 20
 AAB35294
 ID AAB35294 standard; peptide; 18 AA.
 XX
 AC AAB35294;
 XX
 DT 13-NOV-2002 (first entry)
 XX
 XX Synthetic anti-viral human theta defensin peptide HTD-1 SEQ ID NO:27.
 DE
 XX Anti-viral; viral infection; theta-defensin; lipid environment;
 KW amphipathic alpha-helical structure; virucide; anti-HIV; immunisation;
 KW viral growth inhibitor; viral proliferation inhibitor.
 XX
 XX Homo sapiens.
 OS
 OS Synthetic.
 XX
 XX WO2000260468-A2.
 XX
 XX 08-AUG-2002.
 XX
 XX 29-JAN-2002; 2002WO-US002435.
 XX
 XX 30-JAN-2001; 2001US-0265270P.
 XX
 XX 01-AUG-2001; 2001US-0309368P.
 XX

XX (IOWA) UNIV IOWA RES FOUND.
 XX Maury W, Stapleton J, Stinski M, Roller R, Mcoray PB, Tack B;
 XX WPI; 2002-674815/72.
 XX
 XX New method of using a first anti-viral peptide comprising a Theta-
 PT defensin peptide in an amphipathic Alpha-helical structure in a lipid
 PT environment for reducing the infectivity of a virus.
 XX
 XX Disclosure; Page 9; 65pp; English.
 XX
 XX The present invention describes a method (M1) of using a first anti-viral
 CC peptide (I) comprising a theta-defensin peptide in an amphipathic alpha-
 CC helical structure in a lipid environment for reducing the infectivity of
 CC a virus. (I) can have virucide and anti-HIV activities, and can be used
 CC to reduce virus growth, infectivity burden, shed, and development of anti
 CC viral resistance. (I) can be used for inhibiting the growth and
 CC proliferation of a virus and so can be used for; (a) protecting or
 CC treating subject from a viral infection, preventing recurrent viral
 CC infection in a subject harbouring a latent virus, controlling virus
 CC spread within a virally-infected subject (VS), reducing viral burden in a
 CC VS, reducing virus shed from a VS, reducing percentage of VS in a
 CC population regardless of viral infection status, or inducing latency in a
 CC VS; (b) reducing the infectivity of a virus; and (c) rendering virus-
 CC contaminated tissue or fluid sample safe for use, or reducing the number
 CC of infectious virus particles in a population of viruses. (M1) is useful
 CC for reducing the infectivity of a virus in sheep, cattle, horses, swine,
 CC cats, fowl and humans e.g. an enveloped virus infecting humans such as
 CC human immunodeficiency virus (HIV). Preferably, the anti-viral peptide is
 CC administered to a patient who is immunosuppressed or to a subject who is
 CC not infected with the virus, where the first anti-viral peptide is
 CC administered prior to or subsequent to the virus contacting the subject.
 CC The anti-viral peptide is most preferably administered to a subject who
 CC is chronically, latently or acutely infected with the virus. The present
 CC sequence represents a human theta defensin anti-viral peptide, which is
 CC given in the exemplification of the present invention
 XX
 XX Sequence 18 AA;
 SQ
 Query Match 82.1%; Score 92; DB 5; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0013;
 Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GVCRCRCRGVCRCLCRR 18
 Db 1 GICRCICGRCICRCICGR 18
 RESULT 21
 AAE33801
 ID AAE33801 standard; peptide; 18 AA.
 XX
 XX AAE33801;
 AC
 DT 16-APR-2003 (first entry)
 DE
 XX Human retrocyclin peptide.
 XX
 XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
 KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
 KW antibiotic modelling; antimicrobial; human; cyclic.
 XX
 XX Homo sapiens.
 OS
 XX Unidentified.
 XX
 XX Key Location/Qualifiers
 FH Key Location/Qualifiers
 FT Misc-difference 1
 FT /note= "Linked to amino acid at position 18 to form a
 FT cyclic structure"
 FT
 FT Disulfide-bond 3. .16
 FT Disulfide-bond 5. .14
 FT Disulfide-bond 7. .12
 FT Misc-difference 18
 FT /note= "Linked to amino acid at position 1 to form a
 FT cyclic structure"

FT Misc-difference 18
 FT /note= "Linked to amino acid at position 1 to form a
 FT cyclic structure"
 XX
 XX WO200285401-A1.
 PN
 XX
 PD 31-OCT-2002.
 XX
 XX 18-APR-2002; 2002WO-US012353.
 PF
 XX 18-APR-2001; 2001US-0284855P.
 PR
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX Lehrer RI, Waring AJ, Cole AM, Hong TB;
 XX WPI; 2003-103387/09.
 DR
 XX New isolated retrocyclin peptide, useful for preventing retroviral
 PT infections in cells susceptible to bacterial or viral infections or
 PT treating patients having the infections, such as HIV, sexually
 PT transmitted diseases, vaginosis.
 XX
 XX Claim 9; Page 24; 72pp; English.
 PS
 XX The invention relates to novel retrocyclin peptides. Peptides and methods
 CC of the invention are useful for preventing retroviral infections in cells
 CC susceptible to bacterial or viral infections, or treating patients having
 CC infections such as HIV (human immunodeficiency virus), sexually
 CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
 CC retrocyclin-mediated killing is useful for modelling and screening novel
 CC antibiotics. The invention is also useful in gene therapy. The present
 CC sequence is human retrocyclin peptide
 XX
 XX Sequence 18 AA;
 SQ
 Query Match 82.1%; Score 92; DB 6; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0013;
 Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GVCRCRCRGVCRCLCRR 18
 Db 1 GICRCICGRCICRCICGR 18
 RESULT 22
 AAE33863
 ID AAE33863 standard; peptide; 18 AA.
 XX
 XX AAE33863;
 AC
 XX
 DT 16-APR-2003 (first entry)
 DE
 XX Enantio-retrocyclin peptide analogue.
 XX
 XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
 KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
 KW antibiotic modelling; antimicrobial; cyclic.
 XX
 XX Unidentified.
 OS
 XX
 XX Key Location/Qualifiers
 FH Key Location/Qualifiers
 FT Misc-difference 1. .18
 FT /note= "D-form residues"
 FT
 FT Misc-difference 1
 FT /note= "Linked to amino acid at position 18 to form a
 FT cyclic structure"
 FT
 FT Disulfide-bond 3. .16
 FT Disulfide-bond 5. .14
 FT Disulfide-bond 7. .12
 FT Misc-difference 18
 FT /note= "Linked to amino acid at position 1 to form a
 FT cyclic structure"

XX WO200285401-A1.
 XX 31-OCT-2002.
 XX 18-APR-2002; 2002WO-US012353.
 XX 18-APR-2001; 2001US-0284855P.
 XX (REGC) UNIV CALIFORNIA.
 XX Lehrer RI, Waring AJ, Cole AM, Hong TB;
 XX WPI; 2003-103387/09.
 XX New isolated retrocyclin peptide, useful for preventing retroviral
 PT infections in cells susceptible to bacterial or viral infections or
 PT treating patients having the infections, such as HIV, sexually
 PT transmitted diseases, vaginosis.
 XX Disclosure; Page 24; 72pp; English.
 XX The invention relates to novel retrocyclin peptides. Peptides and methods
 CC of the invention are useful for preventing retroviral infections in cells
 CC susceptible to bacterial or viral infections, or treating patients having
 CC infections such as HIV (human immunodeficiency virus), sexually
 CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
 CC retrocyclin-mediated killing is useful for modelling and screening novel
 CC antibiotics. The invention is also useful in gene therapy. The present
 CC sequence is human retrocyclin peptide analogue
 XX Sequence 18 AA;
 SQ Query Match 82.1%; Score 92; DB 6; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0013;
 Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GVCRCCLCRGGVCRCLCRR 18
 Db :|||:|||||:|
 1 GICRCICGICGICICGR 18
 RESULT 23
 ADN08176
 ID ADN08176 standard; peptide; 18 AA.
 XX AC ADN08176;
 XX 15-JUL-2004 (first entry)
 XX Human retrocyclin peptide, RC-100.
 XX retrocyclin; cyclic; primate; retroviral infection;
 KW circular minidefensin; antibacterial; virucide; vaccine; immunotherapy;
 KW microbial; viral; human.
 XX Homo sapiens.
 XX WO2004033479-A2.
 XX 22-APR-2004.
 XX 06-MAY-2003; 2003WO-US014106.
 XX 06-MAY-2002; 2002US-00141645.
 XX (REGC) UNIV CALIFORNIA.
 XX (LEHR/) LEHRER R J.
 XX (WARI/) WARING A J.
 XX (COLE/) COLE A M.
 XX (HONG/) HONG T B.
 XX Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2004-340883/31.
 DR N-PSDB; ADN08193.
 XX New isolated retrocyclin peptides and cyclic polypeptides, useful as
 PT therapeutic and prophylactic agents for treating and preventing microbial
 PT and viral infections.
 XX Claim 9; SEQ ID NO 1; 82pp; English.
 XX The invention relates to a novel isolated retrocyclin peptide. The
 CC invention further provides: a cyclic polypeptide; an isolated nucleic
 CC acid encoding a primate retrocyclin; a method for preventing retroviral
 CC infection in a cell by administering an effective dose of a circular
 CC minidefensin or retrocyclin to the cell; a method for killing microbial
 CC organisms by administering an effective dose of retrocyclin to the
 CC microbial organisms; a method for administering retrocyclin as a
 CC therapeutic agent to a patient with an established microbial or viral
 CC infection; and a method for administering retrocyclin as a prophylactic
 CC agent to prevent a microbial or viral infection in a patient at risk of
 CC developing such infection. The retrocyclin peptide has antibacterial and
 CC virucide activities. The retrocyclin peptide can be used to treat a viral
 CC infection such as HIV-1. The retrocyclin peptide can be used as a vaccine
 CC and can be used in immunotherapy. The peptide and polypeptides are useful
 CC as therapeutic and prophylactic agents for treating and preventing
 CC microbial and viral infections. This sequence represents a retrocyclin
 CC peptide of the invention.
 XX Sequence 18 AA;
 SQ Query Match 82.1%; Score 92; DB 8; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0013;
 Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GVCRCCLCRGGVCRCLCRR 18
 Db :|||:|||||:|
 1 GICRCICGICGICICGR 18
 RESULT 24
 ABP53299
 ID ABP53299 standard; peptide; 18 AA.
 XX AC ABP53299;
 XX 13-NOV-2002 (first entry)
 XX Anti-viral chimeric theta defensin peptide H/RTD-2 SEQ ID NO:32.
 XX Anti-viral; viral infection; theta-defensin; lipid environment;
 KW amphipathic alpha-helical structure; virucide; anti-HIV; immunisation;
 KW viral growth inhibitor; viral proliferation inhibitor.
 XX Homo sapiens.
 OS Macaca mulatta.
 OS Synthetic.
 XX WO200260468-A2.
 XX 08-AUG-2002.
 XX 29-JAN-2002; 2002WO-US002435.
 XX 30-JAN-2001; 2001US-0265270P.
 PR 01-AUG-2001; 2001US-0309368P.
 XX (IOWA) UNIV IOWA RES FOUND.
 XX Maury W, Stapleton J, Stinski M, Roller R, McCray PB, Tack B;
 XX WPI; 2002-674815/72.
 XX New method of using a first anti-viral peptide comprising a Theta-

PR 18-APR-2001; 2001US-0284855P.

XX PA (REGC) UNIV CALIFORNIA.

PI Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2003-103387/09.

XX New isolated retrocyclin peptide, useful for preventing retroviral
PT infections in cells susceptible to bacterial or viral infections or
PT treating patients having the infections, such as HIV, sexually
PT transmitted diseases, vaginosis.

XX Disclosure; Page 24; 72pp; English.

XX The invention relates to novel retrocyclin peptides. Peptides and methods
CC of the invention are useful for preventing retroviral infections in cells
CC susceptible to bacterial or viral infections, or treating patients having
CC infections such as HIV (human immunodeficiency virus), sexually
CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
CC retrocyclin-mediated killing is useful for modelling and screening novel
CC antibiotics. The invention is also useful in gene therapy. The present
CC sequence is human retrocyclin peptide analogue

XX Sequence 18 AA;

Query Match 79.5%; Score 89; DB 6; Length 18;

Best Local Similarity 61.1%; Pred. No. 0.0028;

Matches 11; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVCRCICRGVCRCLCRR 18

DB 1 GICRCICGICRCICGR 18

RESULT 27

AAE33803

ID AAE33803 standard; peptide; 18 AA.

AC AAE33803;

XX 16-APR-2003 (first entry)

DE I6Y retrocyclin peptide analogue.

XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
KW antibiotic modelling; antimicrobial; cyclic.

XX Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "Linked to amino acid at position 18 to form a
FT cyclic structure"

FT Disulfide-bond 3..16

FT Disulfide-bond 5..14

FT Disulfide-bond 7..12

FT Misc-difference 18

FT /note= "Linked to amino acid at position 1 to form a
FT cyclic structure"

XX WO200285401-A1.

PN 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012353.

XX 18-APR-2001; 2001US-0284855P.

XX (REGC) UNIV CALIFORNIA.

PI Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2003-103387/09.

XX New isolated retrocyclin peptide, useful for preventing retroviral
PT infections in cells susceptible to bacterial or viral infections or
PT treating patients having the infections, such as HIV, sexually
PT transmitted diseases, vaginosis.

XX Claim 9; Page 24; 72pp; English.

XX The invention relates to novel retrocyclin peptides. Peptides and methods
CC of the invention are useful for preventing retroviral infections in cells
CC susceptible to bacterial or viral infections, or treating patients having
CC infections such as HIV (human immunodeficiency virus), sexually
CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
CC retrocyclin-mediated killing is useful for modelling and screening novel
CC antibiotics. The invention is also useful in gene therapy. The present
CC sequence is human retrocyclin peptide analogue

XX Sequence 18 AA;

Query Match 79.5%; Score 89; DB 6; Length 18;

Best Local Similarity 66.7%; Pred. No. 0.0028;

Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 GVCRCICRGVCRCLCRR 18

DB 1 GICRCICGICRCICGR 18

RESULT 28

AAE33802

ID AAE33802 standard; peptide; 18 AA.

XX AAE33802;

XX 16-APR-2003 (first entry)

DE R9K retrocyclin peptide analogue.

XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
KW antibiotic modelling; antimicrobial; cyclic.

XX Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "Linked to amino acid at position 18 to form a
FT cyclic structure"

FT Disulfide-bond 3..16

FT Disulfide-bond 5..14

FT Disulfide-bond 7..12

FT Misc-difference 18

FT /note= "Linked to amino acid at position 1 to form a
FT cyclic structure"

XX WO200285401-A1.

PN 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012353.

XX 18-APR-2001; 2001US-0284855P.

XX (REGC) UNIV CALIFORNIA.

PI Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2003-103387/09.

XX New isolated retrocyclin peptide, useful for preventing retroviral
PT infections in cells susceptible to bacterial or viral infections or

PT treating patients having the infections, such as HIV, sexually
 PT transmitted diseases, vaginosis.
 XX
 XX Claim 9; Page 24; 72pp; English.

XX The invention relates to novel retrocyclin peptides. Peptides and methods
 CC of the invention are useful for preventing retroviral infections in cells
 CC susceptible to bacterial or viral infections, or treating patients having
 CC infections such as HIV (human immunodeficiency virus), sexually
 CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
 CC retrocyclin-mediated killing is useful for modelling and screening novel
 CC antibiotics. The invention is also useful in gene therapy. The present
 CC sequence is human retrocyclin peptide analogue

XX Sequence 18 AA;

Query Match 79.5%; Score 89; DB 6; Length 18;

Best Local Similarity 61.1%; Pred. No. 0.0028; 2; Indels 0; Gaps 0;
 Matches 11; Conservative 5; Mismatches 2;

Qy 1 GVCRCLCRGVCRCRCRR 18

Db 1 GICRCICGKICRCICGR 18

RESULT 29

AD035244
 ID AD035244 standard; peptide; 18 AA.

AC AD035244;

XX 15-JUL-2004 (first entry)

DE Rhesus theta defensin analogue peptide 5:3 ARTD-3-NH.

KW Monkey; Rhesus theta defensin; RTD; antimicrobial peptide; antimicrobial;
 KW antiinflammatory; antibacterial; virucide; fungicide; food;
 KW contact lens solution; eye wash solution; inflammatory response;
 KW microbicidal inhibition; microbiostatic growth inhibition; disinfectant;
 KW food preservative; bacterial infection; viral infection;
 KW fungal infection; haemolytic activity.

XX Macaca mulatta.

OS Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 6..15

FT Disulfide-bond 8..13

FT Modified-site 18

FT /note= "Amidated"

XX US2004014669-A1.

XX 22-JAN-2004.

XX 30-APR-2003; 2003US-00427715.

XX 30-APR-2002; 2002US-0377071P.

XX (REGC) UNIV CALIFORNIA.

XX Selsted ME, Tran DQ;

XX WPI; 2004-167945/16.

XX Novel theta defensin analog useful for reducing or inhibiting growth or
 PT survival of a microorganism in an environment such as food or food
 PT product, contact lens solution, or eye wash solution, an inanimate
 PT object.

XX Claim 1; SEQ ID NO 18; 46pp; English.

XX The invention relates to a theta defensin analogue defined by formulae

CC detailed in the claims or appearing as ADO35239-ADO35257. The theta
 CC defensin analogue is useful for reducing or inhibiting growth or survival
 CC of a microorganism in an environment capable of sustaining the growth or
 CC survival of the microorganism and is useful for reducing or inhibiting
 CC growth or survival of a microorganism in an environment such as food or
 CC food product, a solution (e.g., contact lens solution, or eye wash
 CC solution), an inanimate object comprising surface, or a mammal. The
 CC peptides are also useful for decreasing inflammatory response and for
 CC microbicidal inhibition of survival of microorganism as well as
 CC microbiostatic inhibition of growth. Thus the peptides are useful as
 CC therapeutic agents, disinfectants, food preservatives, or medicaments.
 CC The peptides are also useful for treating a patient suffering from
 CC bacterial, viral, fungal or other infection. The theta defensins have
 CC high antimicrobial activity and low haemolytic activity. The present
 CC sequence represents a Rhesus theta defensin analogue peptide.

XX Sequence 18 AA;

Query Match 79.5%; Score 89; DB 8; Length 18;

Best Local Similarity 88.2%; Pred. No. 0.0028;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GVCRCLCRGVCRCRCRR 17

Db 2 GVARCLCRRGVCRCRR 18

RESULT 30

ADN08177
 ID ADN08177 standard; peptide; 18 AA.

AC ADN08177;

XX 15-JUL-2004 (first entry)

XX Human retrocyclin peptide, RC-101.

XX retrocyclin; cyclic; primate; retroviral infection;

KW circular minidefensin; antibacterial; virucide; vaccine; immunotherapy;

KW microbial; viral; human.

XX Homo sapiens.

OS Synthetic.

XX WO2004033479-A2.

XX 22-APR-2004.

XX 06-MAY-2003; 2003WO-US014106.

XX 06-MAY-2002; 2002US-00141645.

XX (REGC) UNIV CALIFORNIA.

XX (LEHR/) LEHRER R I.

XX (WARI/) WARING A J.

XX (COLE/) COLE A M.

XX (HONG/) HONG T B.

XX Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2004-340883/31.

XX New isolated retrocyclin peptides and cyclic polypeptides, useful as
 PT therapeutic and prophylactic agents for treating and preventing microbial
 PT and viral infections.

XX Claim 9; SEQ ID NO 2; 82pp; English.

XX The invention relates to a novel isolated retrocyclin peptide. The
 CC invention further provides a cyclic polypeptide; an isolated nucleic
 CC acid encoding a primate retrocyclin; a method for preventing retroviral
 CC infection in a cell by administering an effective dose of a circular
 CC minidefensin or retrocyclin to the cell; a method for killing microbial

CC as therapeutic and prophylactic agents for treating and preventing
CC microbial and viral infections. This sequence represents a retrocyclin
CC peptide of the invention.

Query Match 79.5%; Score 89; DB 8; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0028;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GVCRLCRRGVCRCLCRR 18
| : | : | : | : | : | : |
pb 1 GICRCICGRGICRCYCGR 18

RESULT 32
ADN08178
IN ADN08178 standard: peptide: 18 AA.

AC ADN08178;

15-JUL-2004 (first entry)

Human retrocyclin peptide, RC-102.

retrocyclin; cyclic; primate; retroviral infection;
circular minidensin; antibacterial; virucide; vaccine; immunotherapy;
microbial; viral; human.

xx Homo sapiens.
OS Synthetic.

XX PN WO2004033479-A2

22-APR-2004 11:22:22

XX
PF 06-MAY-2003: 2003WO-US014106.

06 MAY 2003 2003HS-0014164E

[illegible]

PA (LEHR/) LEHRER R I.

PA (COLE/) COLE A M.

XX

XX
XX

XX New isolated retrocyclin peptides and cyclic polypeptides, useful as
PT therapeutic and prophylactic agents for treating and preventing microbial
PT infection

XXIX

XX The invention relates to a novel isolated retrocyclin peptide. The
CC invention further provides: a cyclic polypeptide; an isolated nucleic
CC acid encoding a primate retrocyclin; a method for preventing retroviral
CC infection in a cell by administering an effective dose of a circular
CC minidensin or retrocyclin to the cell; a method for killing microbial
CC organisms by administering an effective dose of retrocyclin to the
CC microbial organisms; a method for administering retrocyclin as a
CC therapeutic agent to a patient with an established microbial or viral
CC infection; and a method for administering retrocyclin as a prophylactic
CC agent to prevent a microbial or viral infection in a patient at risk of
CC developing such infection. The retrocyclin peptide has antibacterial and

CC virucide activities. The retrocyclin peptide can be used to treat a viral
CC infection such as HIV-1. The retrocyclin peptide can be used as a vaccine
CC and can be used in immunotherapy. The peptide and polypeptides are useful
CC as therapeutic and prophylactic agents for treating and preventing
CC microbial and viral infections. This sequence represents a retrocyclin
CC peptide of the invention.

XX Sequence 18 AA;
SQ Query Match 79.5%; Score 89; DB 8; Length 18;
Best Local Similarity 66.7%; Pred. NO. 0.0028;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVCRCRCRGVCRCLCRR 18
Db 1 GICRCVCGRCICRCICGR 18

RESULT 33
AAE33805
ID AAE33805 standard; peptide; 18 AA.

XX AAE33805;
AC AAE33805;
XX 16-APR-2003 (first entry)

XX 12Y retrocyclin peptide analogue.

XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
KW antibiotic modelling; antimicrobial; cyclic.

XX Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "Linked to amino acid at position 18 to form a

FT cyclic structure"

FT Disulfide-bond 3..16

FT Disulfide-bond 5..14

FT Disulfide-bond 7..12

FT Misc-difference 18

FT /note= "Linked to amino acid at position 1 to form a

FT cyclic structure"

XX WO200285401-A1.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012353.

XX 18-APR-2001; 2001US-0284855P.

XX (REGC) UNIV CALIFORNIA.

XX Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2003-103387/09.

XX New isolated retrocyclin peptide, useful for preventing retroviral
PT infections in cells susceptible to bacterial or viral infections or
PT treating patients having the infections, such as HIV, sexually
PT transmitted diseases, vaginosis.

XX Claim 9; Page 24; 72pp; English.

XX The invention relates to novel retrocyclin peptides. Peptides and methods
CC of the invention are useful for preventing retroviral infections in cells
CC susceptible to bacterial or viral infections, or treating patients having
CC infections such as HIV (human immunodeficiency virus), sexually
CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
CC retrocyclin-mediated killing is useful for modelling and screening novel
CC antibiotics. The invention is also useful in gene therapy. The present

CC sequence is human retrocyclin peptide analogue
XX Sequence 18 AA;
SQ Query Match 78.8%; Score 88; DB 6; Length 18;
Best Local Similarity 66.7%; Pred. NO. 0.0035;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVCRCRCRGVCRCLCRR 18
Db 1 GYCRCICRGICRCICGR 18

RESULT 34
AAE33806
ID AAE33806 standard; peptide; 18 AA.

XX AAE33806;
AC AAE33806;
XX 16-APR-2003 (first entry)

XX 11Y retrocyclin peptide analogue.

XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
KW antibiotic modelling; antimicrobial; cyclic.

XX Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "Linked to amino acid at position 18 to form a

FT cyclic structure"

FT Disulfide-bond 3..16

FT Disulfide-bond 5..14

FT Disulfide-bond 7..12

FT Misc-difference 18

FT /note= "Linked to amino acid at position 1 to form a

FT cyclic structure"

XX WO200285401-A1.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012353.

XX 18-APR-2001; 2001US-0284855P.

XX (REGC) UNIV CALIFORNIA.

XX Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2003-103387/09.

XX New isolated retrocyclin peptide, useful for preventing retroviral
PT infections in cells susceptible to bacterial or viral infections or
PT treating patients having the infections, such as HIV, sexually
PT transmitted diseases, vaginosis.

XX Claim 9; Page 24; 72pp; English.

XX The invention relates to novel retrocyclin peptides. Peptides and methods
CC of the invention are useful for preventing retroviral infections in cells
CC susceptible to bacterial or viral infections, or treating patients having
CC infections such as HIV (human immunodeficiency virus), sexually
CC transmitted diseases, bacterial vaginosis or ophthalmic infections. The
CC retrocyclin-mediated killing is useful for modelling and screening novel
CC antibiotics. The invention is also useful in gene therapy. The present
CC sequence is human retrocyclin peptide analogue

XX Sequence 18 AA;
SQ Query Match 78.8%; Score 88; DB 6; Length 18;

RESULT 37
 AAB35046
 ID AAB35046 standard; peptide; 18 AA.
 XX AC AAB35046;
 XX DT 27-MAR-2001 (first entry)
 XX DE Theta defensin SEQ ID NO: 30.
 XX KW Theta defensin; antimicrobial; cyclic; bacterium; fungus; protozoan;
 XX KW virus; helminth; disinfectant; food preservative; analogue.
 XX OS Unidentified.
 XX PN WO200068265-A1.
 XX PD 16-NOV-2000.
 XX PF 10-MAY-2000; 2000WO-US012842.
 XX PR 10-MAY-1999; 99US-00309487.
 XX PA (REGC) UNIV CALIFORNIA.
 XX PI Selsted ME, Tang Y, Yuan J, Ouellette AJ;
 XX WPI; 2001-031853/04.
 XX Novel theta defensin peptide with antimicrobial activity against
 PT bacteria, yeast, fungi, protozoa and viruses.
 PS Claim 15; Fig 16; 110pp; English.
 XX The present invention provides theta defensin peptides and analogues
 CC which have antimicrobial activity. They can be used in the treatment of
 CC bacterial, viral, fungal, protozoan and helminthic infections, in
 CC disinfectants and as food preservatives
 XX Sequence 18 AA;
 SQ
 Query Match 76.8%; Score 86; DB 4; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0056;
 Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GVCRCLCRRGVCRCLCRR 18
 DB 1 GFCRCICTRGFCRCICTR 18
 RESULT 38
 ABP53295
 ID ABP53295 standard; peptide; 18 AA.
 XX AC ABP53295;
 XX DT 13-NOV-2002 (first entry)
 XX DE Anti-viral theta defensin peptide RTD-2 SEQ ID NO:28.
 XX KW Anti-viral; viral infection; theta-defensin; lipid environment;
 XX KW amphipathic alpha-helical structure; virucide; anti-HIV; immunisation;
 XX KW viral growth inhibitor; viral proliferation inhibitor.
 XX OS Macaca mulatta.
 XX OS Synthetic.
 XX PN WO200260468-A2.
 XX PD 08-AUG-2002.
 XX
 29-JAN-2002; 2002WO-US002435.
 XX 30-JAN-2001; 2001US-0265270P.
 PR 01-AUG-2001; 2001US-0309368P.
 XX (IOWA) UNIV IOWA RES FOUND.
 XX Maury W, Stapleton J, Stinski M, Roller R, Mccray PB, Tack B;
 XX WPI; 2002-674815/72.
 XX New method of using a first anti-viral peptide comprising a Theta-
 PT defensin peptide in an amphipathic Alpha-helical structure in a lipid
 PT environment for reducing the infectivity of a virus.
 XX Disclosure; Page 10; 65pp; English.
 XX The present invention describes a method (M1) of using a first anti-viral
 CC peptide (I) comprising a theta-defensin peptide in an amphipathic alpha-
 CC helical structure in a lipid environment for reducing the infectivity of
 CC a virus. (I) can have virucide and anti-HIV activities, and can be used
 CC to reduce virus growth, infectivity burden, shed, and development of anti-
 CC -viral resistance. (I) can be used for inhibiting the growth and
 CC proliferation of a virus and so can be used for: (a) protecting or
 CC treating subject from a viral infection, preventing recurrent viral
 CC infection in a subject harbouring a latent virus, controlling virus
 CC spread within a virally-infected subject (VS), reducing viral burden in a
 CC VS, reducing virus shed from a VS, reducing percentage of VS in a
 CC population regardless of viral infection status, or inducing latency in a
 CC VS; (b) reducing the infectivity of a virus; and (c) rendering virus-
 CC contaminated tissue or fluid sample safe for use, or reducing the number
 CC of infectious virus particles in a population of viruses. (M1) is useful
 CC for reducing the infectivity of a virus in sheep, cattle, horses, swine,
 CC cats, fowl and humans e.g. an enveloped virus infecting humans such as
 CC human immunodeficiency virus (HIV). Preferably, the anti-viral peptide is
 CC administered to a patient who is immunosuppressed or to a subject who is
 CC not infected with the virus, where the first anti-viral peptide is
 CC administered prior to or subsequent to the virus contacting the subject.
 CC The anti-viral peptide is most preferably administered to a subject who
 CC is chronically, latently or acutely infected with the virus. The present
 CC sequence represents a rhesus monkey theta defensin anti-viral peptide,
 CC which is given in the exemplification of the present invention
 XX Sequence 18 AA;
 SQ
 Query Match 76.8%; Score 86; DB 5; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0056;
 Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GVCRCLCRRGVCRCLCRR 18
 DB 1 GFCRCICTRGFCRCICTR 18
 RESULT 39
 AAE33865
 ID AAE33865 standard; peptide; 18 AA.
 XX AC AAE33865;
 XX DT 16-APR-2003 (first entry)
 XX DE RC-101/103 retrocyclin peptide analogue.
 XX KW Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
 XX KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
 XX KW antibiotic modelling; antimicrobial; cyclic.
 XX OS Unidentified.
 XX Key Location/Qualifiers
 PH Misc-difference 1 /note= "Linked to amino acid at position 18 to form a
 FT

